

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 05:24:49 ; Search time 25 Seconds

(without alignments)
56.667 Million cell updates/sec

Title: US-09-727-892a-99

Perfect score: 58

Sequence: 1 MERKRYVLYLCDEIKGHF.....YEVNLPFKKAVITIEYKEI 58

Scoring table: OLIGO

Searched: 231628 seqs, 24425594 residues

Word size: 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: Issued Patents_AA:*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUTS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	10.3	98	US-08-851-843A-12	Sequence 12, Appl
2	6	10.3	98	US-08-854-050-12	Sequence 12, Appl
3	6	10.3	98	US-09-430-323-12	Sequence 12, Appl
4	6	10.3	99	US-08-974-549A-194	Sequence 194, App
5	6	10.3	109	US-09-187-859-8	Sequence 8, Appl
6	6	10.3	378	US-08-401-068-14	Sequence 14, Appl
7	6	10.3	378	US-08-846-338-14	Sequence 14, Appl
8	6	10.3	380	US-08-472-659-34	Sequence 34, Appl
9	6	10.3	380	US-08-474-661-34	Sequence 34, Appl
10	6	10.3	380	US-08-611-977-34	Sequence 34, Appl
11	6	10.3	422	US-08-680-726A-68	Sequence 68, Appl
12	6	10.3	422	US-09-092-409-68	Sequence 68, Appl
13	6	10.3	450	US-09-306-593-13	Sequence 13, Appl
14	6	10.3	485	US-09-058-260-18	Sequence 18, Appl
15	6	10.3	501	US-08-781-802-4	Sequence 4, Appl
16	6	10.3	501	US-08-781-802-6	Sequence 6, Appl
17	6	10.3	501	US-08-694-078-4	Sequence 4, Appl
18	6	10.3	501	US-08-694-078-6	Sequence 6, Appl
19	6	10.3	501	US-09-058-260-4	Sequence 4, Appl
20	6	10.3	501	US-09-058-260-6	Sequence 6, Appl
21	6	10.3	501	US-09-058-260-10	Sequence 10, Appl
22	6	10.3	501	US-09-058-260-14	Sequence 14, Appl
23	6	10.3	501	US-09-058-260-20	Sequence 20, Appl
24	6	10.3	501	US-09-058-260-22	Sequence 22, Appl
25	6	10.3	501	US-09-058-260-24	Sequence 24, Appl
26	6	10.3	513	US-08-676-166A-5	Sequence 5, Appl
27	6	10.3	514	US-09-254-733-2	Sequence 2, Appl

28	6	10.3	550	1	US-08-674-168-29	Sequence 29, Appl
29	6	10.3	550	3	US-08-985-908-19	Sequence 19, Appl
30	6	10.3	550	3	US-08-852-730-4	Sequence 4, Appl
31	6	10.3	550	4	US-08-985-916-11	Sequence 11, Appl
32	6	10.3	699	3	US-08-851-843A-52	Sequence 52, Appl
33	6	10.3	699	4	US-08-974-549A-188	Sequence 188, App
34	6	10.3	699	4	US-08-854-050-52	Sequence 52, Appl
35	6	10.3	699	4	US-09-430-323-52	Sequence 52, Appl
36	6	10.3	719	3	US-08-851-843A-7	Sequence 7, Appl
37	6	10.3	719	4	US-08-974-549A-219	Sequence 219, App
38	6	10.3	719	4	US-08-854-050-7	Sequence 7, Appl
39	6	10.3	719	4	US-09-430-323-7	Sequence 7, Appl
40	5	8.6	5	2	US-08-472-659-7	Sequence 7, Appl
41	5	8.6	5	2	US-08-474-661-7	Sequence 7, Appl
42	5	8.6	5	2	US-08-611-977-7	Sequence 7, Appl
43	5	8.6	9	2	US-08-340-283-146	Sequence 146, Appl
44	5	8.6	14	1	US-08-383-753-24	Sequence 24, Appl
45	5	8.6	14	2	US-08-586-772-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-08-851-843A-12
Sequence 12, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morlin, Gregg B.
APPLICANT: Hartley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: NO. 6093809e1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-00293005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-12

Query Match 10.3%; Score 6; DB 3; Length 98;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 DEIKGH 18
14 DEIKGH 19

RESULT 2
US-08-854-050-12
Sequence 12, Application US/08854050
Patent No. 6261836

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6261836el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-00293005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids

TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-12

Query Match 10.3%; Score 6; DB 4; Length 98;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 DEIKGH 18
14 DEIKGH 19

RESULT 3
US-09-430-323-12
Sequence 12, Application US/09430323
Patent No. 6309867

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-OCT-1999
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-00293005

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid

STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-430-323-12

Query Match 10.3%; Score 6; DB 4; Length 98;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 DEIKGH 18
Db 14 DEIKGH 19

RESULT 4

US-08-974-549A-194
; Sequence 194, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cecch, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 194:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

MOLECULE TYPE: peptide
US-08-974-549A-194

Query Match 10.3%; Score 6; DB 4; Length 99;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 DEIKGH 18
Db 14 DEIKGH 19

RESULT 5

US-09-187-859-8
; Sequence 8, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; FILE REFERENCE: 100086.407C1
; CURRENT APPLICATION NUMBER: US/09/187,859A
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-187-859-8

Query Match 10.3%; Score 6; DB 4; Length 109;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 AKVYVS 37
Db 36 AKVYVS 41

RESULT 6
US-08-401-068-14
; Sequence 14, Application US/08401068
; Patent No. 5859335
; GENERAL INFORMATION:
; APPLICANT: Patton, David

; TITLE OF INVENTION: Enhanced Biotin Biosynthesis in Plant Tissue
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/401,068
; FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/351,970
FILING DATE: 08-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 378 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-401-068-14

Query Match 10.3%; Score 6; DB 2; Length 378;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 EYKEI 58
|||||
DB 158 EYKEI 163

RESULT 7
US-08-846-338-14
Sequence 14, Application US/08846338
Patent No. 5869719
GENERAL INFORMATION:
APPLICANT: Patton, David
TITLE OF INVENTION: Transgenic Plants Having Increased Biotin Content
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5869719artis Corporation
STREET: 520 White Plains Road, P.O. Box 2005
City: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/846,338
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 378 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-846-338-14

Query Match 10.3%; Score 6; DB 2; Length 378;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 EYKEI 58
|||||
DB 158 EYKEI 163

RESULT 8
US-08-472-659-34
Sequence 34, Application US/08472659
Patent No. 5831030
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROOKA, No. 5831030uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5831030uhitro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5831030omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
City: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
APPLICANT INFORMATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-659-34

Query Match 10.3%; Score 6; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 LYDAKV 34
|||||
DB 121 LYDAKV 126

RESULT 9
US-08-474-661-34
Sequence 34, Application US/08474661
Patent No. 5874253
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki

APPLICANT: TSURUOKA, No. 5874253uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5874253uniro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,661
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANER
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-474-661-34

Query Match 10.3%; Score 6; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 LYDAKV 34
Db 121 LYDAKV 126

RESULT 10
US-08-611-977-34
Sequence 34, Application US/08611977
Patent No. 5972886
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5972886uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5972886uniro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,977
FILING DATE: 06-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-611-977-34

Query Match 10.3%; Score 6; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 LYDAKV 34
Db 121 LYDAKV 126

RESULT 11
US-08-680-726A-68
Sequence 68, Application US/08680726A
Patent No. 5804197
GENERAL INFORMATION:
APPLICANT: Haanes, Elizabeth J.
APPLICANT: Frank, Rexann S.
TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUSES
NUMBER OF SEQUENCES: 92
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/680,726A
FILING DATE: 12-JUL-1996

CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-46-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 422 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-680-726A-68

Query Match 10.3%; Score 6; DB 1; Length 422;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 EYNLFT 45
|||||
DB 398 EYNLFT 403

RESULT 12
US-09-092-409-68
Sequence 68, Application US/09092409
Patent No. 6159478
GENERAL INFORMATION:
APPLICANT: Haanes, Elizabeth J.
APPLICANT: Frank, Rexann S.
TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUSES
NUMBER OF SEQUENCES: 92
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/092,409
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/680,726
FILING DATE: 12-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-46-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 422 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-092-409-68

Query Match 10.3%; Score 6; DB 4; Length 422;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 EYNLFT 45

|||||
DB 398 EYNLFT 403

RESULT 13
US-09-306-593-13
Sequence 13, Application US/09306593
Patent No. 6184018
GENERAL INFORMATION:
APPLICANT: Li, Xin-Liang
APPLICANT: Ljungdahl, Lars G.
APPLICANT: Chen, Huizhong
APPLICANT: Ximenes, Eduardo A.
TITLE OF INVENTION: Beta-glucosidase Coding Sequences and Protein from
FILE REFERENCE: 31-98us
CURRENT APPLICATION NUMBER: US/09/306,593
CURRENT FILING DATE: 1999-05-06
EARLIER APPLICATION NUMBER: US 60/084,494
EARLIER FILING DATE: 1998-05-06
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 13
LENGTH: 450
TYPE: PRT
ORGANISM: Thermoaerobacter brockii
US-09-306-593-13

Query Match 10.3%; Score 6; DB 4; Length 450;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 52 IEYIKE 57
|||||
DB 376 IEYIKE 381

RESULT 14
US-09-058-260-18
Sequence 18, Application US/09058260B
Patent No. 6218167
GENERAL INFORMATION:
APPLICANT: Allen, Larry
APPLICANT: Aikens, John
APPLICANT: Fongstein, Michael
APPLICANT: Vonstein, Veronika
APPLICANT: Demirjian, David
APPLICANT: Casadaban, Malcolm
TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
FILE REFERENCE: 95-963-H
CURRENT APPLICATION NUMBER: US/09/058,260B
CURRENT FILING DATE: 1999-04-10
EARLIER APPLICATION NUMBER: 60/001,995
EARLIER FILING DATE: 1996-08-07
EARLIER APPLICATION NUMBER: 60/009,704
EARLIER FILING DATE: 1996-01-11
EARLIER APPLICATION NUMBER: 60/019,580
EARLIER FILING DATE: 1996-06-12
EARLIER APPLICATION NUMBER: 08/694,078
EARLIER FILING DATE: 1996-08-08
EARLIER APPLICATION NUMBER: 08/781,802
EARLIER FILING DATE: 1997-01-10
EARLIER APPLICATION NUMBER: 08/827,810
EARLIER FILING DATE: 1997-04-11
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 18
LENGTH: 485
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: cloned esterase
OTHER INFORMATION: gene from bacteria E008

US-09-058-260-18

Query Match 10.3%; Score 6; DB 4; Length 485;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EYNLFT 45
|||||
DB 296 EYNLFT 301

RESULT 15

US-08-781-802-4
; Sequence 4, Application US/08781802
; Patent No. 5969121
; GENERAL INFORMATION:
; APPLICANT: ALLEN, Larry
; APPLICANT: AIKENS, John
; APPLICANT: FONSTEIN, Michael
; APPLICANT: VONSTEIN, Veronika
; APPLICANT: DEMIRJIAN, David
; APPLICANT: CASADABAN, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,802
; FILING DATE: 10-JAN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,078
; FILING DATE: 07-AUG-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/019,580
; FILING DATE: 12-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,704
; FILING DATE: 11-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/001,995
; FILING DATE: 01-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Chao, Mark
; REGISTRATION NUMBER: 37,293
; REFERENCE/DOCKET NUMBER: 95,963-E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 501 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-781-802-4

Query Match 10.3%; Score 6; DB 2; Length 501;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EYNLFT 45

DB 312 EYNLFT 317
|||||

RESULT 16
US-08-781-802-6
; Sequence 6, Application US/08781802
; Patent No. 5969121
; GENERAL INFORMATION:
; APPLICANT: ALLEN, Larry
; APPLICANT: AIKENS, John
; APPLICANT: FONSTEIN, Michael
; APPLICANT: VONSTEIN, Veronika
; APPLICANT: DEMIRJIAN, David
; APPLICANT: CASADABAN, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,802
; FILING DATE: 10-JAN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,078
; FILING DATE: 07-AUG-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/019,580
; FILING DATE: 12-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,704
; FILING DATE: 11-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/001,995
; FILING DATE: 01-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Chao, Mark
; REGISTRATION NUMBER: 37,293
; REFERENCE/DOCKET NUMBER: 95,963-E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 501 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-781-802-6

Query Match 10.3%; Score 6; DB 2; Length 501;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EYNLFT 45
|||||
DB 312 EYNLFT 317

RESULT 17
US-08-694-078-4
; Sequence 4, Application US/08694078

Patent No. 6218163
GENERAL INFORMATION:
APPLICANT: ALLEN, Larry
APPLICANT: AIKENS, John
APPLICANT: FONSEIN, Michael
APPLICANT: VONSTEIN, Veronika
APPLICANT: DEMIRJIAN, David
APPLICANT: CASADABAN, Malcolm
TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff, Ltd.
STREET: 300 S. Wacker Drive 7th Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/694,078
FILING DATE: 07-AUG-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/019,580
FILING DATE: 12-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/009,704
FILING DATE: 10-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/001,995
FILING DATE: 07-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Chao, Mark
REGISTRATION NUMBER: 37,293
REFERENCE/DOCKET NUMBER: 95,963-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-694-078-4

Query Match 10.3%; Score 6; DB 4; Length 501;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EYNLEF 45
|||||
DB 312 EYNLEF 317

RESULT 18
US-08-694-078-6
Sequence 6, Application US/08694078
Patent No. 6218163
GENERAL INFORMATION:
APPLICANT: ALLEN, Larry
APPLICANT: AIKENS, John
APPLICANT: FONSEIN, Michael
APPLICANT: VONSTEIN, Veronika
APPLICANT: DEMIRJIAN, David
APPLICANT: CASADABAN, Malcolm
TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:

ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff, Ltd.
STREET: 300 S. Wacker Drive 7th Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/694,078
FILING DATE: 07-AUG-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/019,580
FILING DATE: 12-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/009,704
FILING DATE: 10-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/001,995
FILING DATE: 07-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Chao, Mark
REGISTRATION NUMBER: 37,293
REFERENCE/DOCKET NUMBER: 95,963-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-694-078-6

Query Match 10.3%; Score 6; DB 4; Length 501;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EYNLEF 45
|||||
DB 312 EYNLEF 317

RESULT 19
US-09-058-260-4
Sequence 4, Application US/09058260B
Patent No. 6218167
GENERAL INFORMATION:
APPLICANT: ALLEN, Larry
APPLICANT: AIKENS, John
APPLICANT: FONSEIN, Michael
APPLICANT: VONSTEIN, Veronika
APPLICANT: DEMIRJIAN, David
APPLICANT: CASADABAN, Malcolm
TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
FILE REFERENCE: 95-963-H
CURRENT APPLICATION NUMBER: US/09/058,260B
FILING DATE: 1999-04-10
EARLIER APPLICATION NUMBER: 60/001,995
FILING DATE: 1996-08-07
EARLIER APPLICATION NUMBER: 60/009,704
FILING DATE: 1996-01-11
EARLIER APPLICATION NUMBER: 60/019,580
FILING DATE: 1996-06-12
EARLIER APPLICATION NUMBER: 08/694,078
FILING DATE: 1996-08-08
EARLIER APPLICATION NUMBER: 08/781,802
FILING DATE: 1997-01-10

```

; EARLIER APPLICATION NUMBER: 08/827,810
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cloned esterase
; OTHER INFORMATION: gene from bacteria E009
US-09-058-260-4

Query Match
Best Local Similarity 10.3%; Score 6; DB 4; Length 501;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EYNLFT 45
DB 312 EYNLFT 317

RESULT 20
US-09-058-260-6
; Sequence 6, Application US/09058260B
; Patent No. 6218167
; GENERAL INFORMATION:
; APPLICANT: Allen, Larry
; APPLICANT: Aikens, John
; APPLICANT: Fonstein, Michael
; APPLICANT: Vonstein, Veronika
; APPLICANT: Demirjian, David
; APPLICANT: Casadaban, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; FILE REFERENCE: 95-963-H
; CURRENT APPLICATION NUMBER: US/09/058,260B
; EARLIER FILING DATE: 1999-04-10
; EARLIER APPLICATION NUMBER: 60/001,995
; EARLIER FILING DATE: 1996-08-07
; EARLIER APPLICATION NUMBER: 60/009,704
; EARLIER FILING DATE: 1996-01-11
; EARLIER APPLICATION NUMBER: 60/019,580
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: 08/694,078
; EARLIER FILING DATE: 1996-08-08
; EARLIER APPLICATION NUMBER: 08/781,802
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 08/827,810
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 6
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cloned esterase
; OTHER INFORMATION: gene from bacteria E011
US-09-058-260-6

Query Match
Best Local Similarity 10.3%; Score 6; DB 4; Length 501;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EYNLFT 45
DB 312 EYNLFT 317

RESULT 21
US-09-058-260-14
; Sequence 14, Application US/09058260B
; Patent No. 6218167
; GENERAL INFORMATION:
; APPLICANT: Allen, Larry
; APPLICANT: Aikens, John
; APPLICANT: Fonstein, Michael
; APPLICANT: Vonstein, Veronika
; APPLICANT: Demirjian, David
; APPLICANT: Casadaban, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; FILE REFERENCE: 95-963-H
; CURRENT APPLICATION NUMBER: US/09/058,260B
; EARLIER FILING DATE: 1999-04-10
; EARLIER APPLICATION NUMBER: 60/001,995
; EARLIER FILING DATE: 1996-08-07
; EARLIER APPLICATION NUMBER: 60/009,704
; EARLIER FILING DATE: 1996-01-11
; EARLIER APPLICATION NUMBER: 60/019,580
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: 08/694,078
; EARLIER FILING DATE: 1996-08-08
; EARLIER APPLICATION NUMBER: 08/781,802
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 08/827,810
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 20

Query Match
Best Local Similarity 10.3%; Score 6; DB 4; Length 501;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EYNLFT 45
DB 312 EYNLFT 317

RESULT 22
US-09-058-260-20
; Sequence 20, Application US/09058260B
; Patent No. 6218167
; GENERAL INFORMATION:
; APPLICANT: Allen, Larry
; APPLICANT: Aikens, John
; APPLICANT: Fonstein, Michael
; APPLICANT: Vonstein, Veronika
; APPLICANT: Demirjian, David
; APPLICANT: Casadaban, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; FILE REFERENCE: 95-963-H
; CURRENT APPLICATION NUMBER: US/09/058,260B
; EARLIER FILING DATE: 1999-04-10
; EARLIER APPLICATION NUMBER: 60/001,995
; EARLIER FILING DATE: 1996-08-07
; EARLIER APPLICATION NUMBER: 60/009,704
; EARLIER FILING DATE: 1996-01-11
; EARLIER APPLICATION NUMBER: 60/019,580
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: 08/694,078
; EARLIER FILING DATE: 1996-08-08
; EARLIER APPLICATION NUMBER: 08/781,802
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 08/827,810
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 20
```

LENGTH: 501
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:cloned esterase
US-09-058-260-20

Query Match 10.3%; Score 6; DB 4; Length 501;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 EYNLFT 45
|||||
DB 312 EYNLFT 317

RESULT 23
US-09-058-260-22
Sequence 22, Application US/09058260B
Patent No. 6218167
GENERAL INFORMATION:
APPLICANT: Allen, Larry
APPLICANT: Aikens, John
APPLICANT: Vonstein, Michael
APPLICANT: Demirjian, Veronika
APPLICANT: Casadaban, Malcolm
TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
FILE REFERENCE: 95-963-H
CURRENT APPLICATION NUMBER: US/09/058, 260B
EARLIER FILING DATE: 1999-04-10
EARLIER APPLICATION NUMBER: 60/001,995
EARLIER FILING DATE: 1996-08-07
EARLIER APPLICATION NUMBER: 60/009,704
EARLIER FILING DATE: 1996-01-11
EARLIER APPLICATION NUMBER: 60/019,580
EARLIER FILING DATE: 1996-06-12
EARLIER APPLICATION NUMBER: 08/694,078
EARLIER FILING DATE: 1996-08-08
EARLIER APPLICATION NUMBER: 08/781,802
EARLIER FILING DATE: 1997-01-10
EARLIER APPLICATION NUMBER: 08/827,810
EARLIER FILING DATE: 1997-04-11
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 22
LENGTH: 501
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:cloned esterase
US-09-058-260-22

Query Match 10.3%; Score 6; DB 4; Length 501;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 EYNLFT 45
|||||
DB 312 EYNLFT 317

RESULT 24
US-09-058-260-24
Sequence 24, Application US/09058260B
Patent No. 6218167
GENERAL INFORMATION:
APPLICANT: Allen, Larry
APPLICANT: Aikens, John
APPLICANT: Vonstein, Michael
APPLICANT: Demirjian, Veronika

APPLICANT: Demirjian, David
APPLICANT: Casadaban, Malcolm
TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
FILE REFERENCE: 95-963-H
CURRENT APPLICATION NUMBER: US/09/058,260B
CURRENT FILING DATE: 1999-04-10
EARLIER APPLICATION NUMBER: 60/001,995
EARLIER FILING DATE: 1996-08-07
EARLIER APPLICATION NUMBER: 60/009,704
EARLIER FILING DATE: 1996-01-11
EARLIER APPLICATION NUMBER: 60/019,580
EARLIER FILING DATE: 1996-06-12
EARLIER APPLICATION NUMBER: 08/694,078
EARLIER FILING DATE: 1996-08-08
EARLIER APPLICATION NUMBER: 08/781,802
EARLIER FILING DATE: 1997-01-10
EARLIER APPLICATION NUMBER: 08/827,810
EARLIER FILING DATE: 1997-04-11
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 24
LENGTH: 501
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:cloned esterase
US-09-058-260-24

Query Match 10.3%; Score 6; DB 4; Length 501;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 EYNLFT 45
|||||
DB 312 EYNLFT 317

RESULT 25
US-09-058-260-32
Sequence 32, Application US/09058260B
Patent No. 6218167
GENERAL INFORMATION:
APPLICANT: Allen, Larry
APPLICANT: Aikens, John
APPLICANT: Vonstein, Michael
APPLICANT: Demirjian, Veronika
APPLICANT: Casadaban, Malcolm
TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
FILE REFERENCE: 95-963-H
CURRENT APPLICATION NUMBER: US/09/058,260B
CURRENT FILING DATE: 1999-04-10
EARLIER APPLICATION NUMBER: 60/001,995
EARLIER FILING DATE: 1996-08-07
EARLIER APPLICATION NUMBER: 60/009,704
EARLIER FILING DATE: 1996-01-11
EARLIER APPLICATION NUMBER: 60/019,580
EARLIER FILING DATE: 1996-06-12
EARLIER APPLICATION NUMBER: 08/694,078
EARLIER FILING DATE: 1996-08-08
EARLIER APPLICATION NUMBER: 08/781,802
EARLIER FILING DATE: 1997-01-10
EARLIER APPLICATION NUMBER: 08/827,810
EARLIER FILING DATE: 1997-04-11
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 32
LENGTH: 501
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:cloned esterase

OTHER INFORMATION: gene from bacteria E027
US-09-058-260-32

Query Match 10.3%; Score 6; DB 4; Length 501;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EYNLEF 45
|||||
DB 312 EYNLEF 317

RESULT 26
US-08-676-166A-5
; Sequence 5, Application US/08676166A
; Patent No. 5955270
; GENERAL INFORMATION:
; APPLICANT: Radford, Alan
; APPLICANT: Parish, John H.
; TITLE OF INVENTION: EXPLOITATION OF THE CELLULOSE COMPLEX OF
; TITLE OF INVENTION: NEUROSPORA
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/676,166A
; FILING DATE: 15-JUL-1996
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1321-1-002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684

INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: T. viride
; US-08-676-166A-5

Query Match 10.3%; Score 6; DB 2; Length 513;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 AKVYVS 37
|||||
DB 431 AKVYVS 436

RESULT 27
US-09-254-733-2
; Sequence 2, Application US/09254733
; Patent No. 6277596
; GENERAL INFORMATION:
; APPLICANT: WATANABE, MANABU

APPLICANT: MORIYA, TATSUKI
; APPLICANT: AOYAGI, KAORI
; APPLICANT: SUMIDA, NAOMI
; TITLE OF INVENTION: REGULATORY SEQUENCE OF CELLULOSE CBHI GENES ORIGINATING
; TITLE OF INVENTION: IN TRICHODERMA VIRIDE AND SYSTEM FOR MASS-PRODUCING
; TITLE OF INVENTION: PROTEINS OR PEPTIDES THEREWITH
; FILE REFERENCE: 99-0266*/LC(WMC)/00144
; CURRENT APPLICATION NUMBER: US/09/254,733
; CURRENT FILING DATE: 1999-05-07
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 514
; TYPE: PRT
; ORGANISM: TRICHODERMA VIRIDE MC300-1
; US-09-254-733-2

Query Match 10.3%; Score 6; DB 4; Length 514;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 AKVYVS 37
|||||
DB 431 AKVYVS 436

RESULT 28
US-08-674-168-29
; Sequence 29, Application US/08674168
; Patent No. 5804414
; GENERAL INFORMATION:
; APPLICANT: MORIYA, MIKA
; APPLICANT: MATSUI, HIROSHI
; APPLICANT: YOKOZAKI, KENZO
; APPLICANT: HIRANO, SEIKO
; APPLICANT: HAYAKAWA, ATSUSHI
; APPLICANT: IZUI, MASAKO

APPLICANT: SUGIMOTO, MASAKAZU
; TITLE OF INVENTION: METHOD OF AMPLIFYING GENE USING
; TITLE OF INVENTION: ARTIFICIAL TRANSPOSON
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P.C.
; STREET: 1755 JEFFERSON DAVIS HIGHWAY, SUITE # 400
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/674,168
; FILING DATE: 01-JUL-1996
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-166541
; FILING DATE: 30-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 10-810-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-674-168-29

Query Match 10.3%; Score 6; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 EYIKI 58
|||||
DB 198 EYIKI 203

RESULT 29
US-08-985-908-19
Sequence 19, Application US/08985908

PATENT No. 6004773

GENERAL INFORMATION:

APPLICANT: MASAYUKI ARAKI, MASAKAZU SUGIMOTO, YASUHIKO YOSHIHARA, AND TSUYOSHI NA

TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.

STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR

CITY: ARLINGTON

STATE: VA

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/985,908

FILING DATE: 05-DEC-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 8-325659

FILING DATE: 05-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: NORMAN F. OBLON

REGISTRATION NUMBER: 24,618

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 550 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-985-908-19

QY 53 EYIKI 58
|||||
DB 198 EYIKI 203

RESULT 30
US-08-852-730-4
Sequence 4, Application US/08852730

PATENT No. 6090597

GENERAL INFORMATION:

APPLICANT: SEIKO HIRANO, MASAKAZU SUGIMOTO, EIICHI NAKANO,

APPLICANT: MASAKO IZUI, ATSUSHI HAYAKAWA, YASUHIKO YOSHIHARA, AND TSUYOSHI

APPLICANT: NAKAMATSU

TITLE OF INVENTION: METHOD OF PRODUCING L-LYSINE

NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER AND NEUSTADT, P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/852,730

FILING DATE: 05-07-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 8-142812

FILING DATE: 05-JUN-1996

ATTORNEY/AGENT INFORMATION:

NAME: NORMAN F. OBLON

REGISTRATION NUMBER: 24,618

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 550 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-852-730-4

QY 53 EYIKI 58
|||||
DB 198 EYIKI 203

RESULT 31
US-08-985-916-11
Sequence 11, Application US/08985916

PATENT No. 6221636

GENERAL INFORMATION:

APPLICANT: ATSUSHI HAYAKAWA, MASAKAZU SUGIMOTO, YASUHIKO YOSHIHARA, AND TSUYOSHI

TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.

STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR

CITY: ARLINGTON

STATE: VA

ZIP: 22152

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/985,916

FILING DATE: 05-DEC-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 8-325658

FILING DATE: 05-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: NORMAN F. OBLON

REGISTRATION NUMBER: 24,618

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 550 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-985-916-11

Query Match 10.3%; Score 6; DB 4; Length 550;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 53 EYKEI 58
|||||
DB 198 EYKEI 203

RESULT 32
US-08-851-843A-52
Sequence 52, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin H.
TITLE OF INVENTION: No. 6093809el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-00293005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 699 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-08-851-843A-52

Query Match 10.3%; Score 6; DB 3; Length 699;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 DEIKGH 18
|||||
DB 486 DEIKGH 491

RESULT 33
US-08-974-549A-188
Sequence 188, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 188:
SEQUENCE CHARACTERISTICS:
LENGTH: 699 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-188

Query Match 10.3%; Score 6; DB 4; Length 699;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 DEIKGH 18
|||||
Db 486 DEIKGH 491

RESULT 34
US-08-854-050-52
Sequence 52, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cecch, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morlin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6261836el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-0029300S
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 699 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-854-050-52

Query Match 10.3%; Score 6; DB 4; Length 699;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 DEIKGH 18
|||||
Db 486 DEIKGH 491

RESULT 35
US-09-430-323-52
Sequence 52, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cecch, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morlin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-OCT-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-0029300S
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 699 amino acids

TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-09-430-323-52

Query Match 10.3%; Score 6; DB 4; Length 699;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 DEIKGH 18
|||||
Db 486 DEIKGH 491

RESULT 36
US-08-851-843A-7

Sequence 7, Application US/08851843A
Patent No. 6093809

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin

APPLICANT: Andrews, William H.

TITLE OF INVENTION: No. 6093809e1 Telomerase

NUMBER OF SEQUENCES: 225

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/851,843A

FILING DATE: 06-MAY-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002930US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 719 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: protein

US-08-851-843A-7

Query Match 10.3%; Score 6; DB 3; Length 719;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 DEIKGH 18
|||||
Db 504 DEIKGH 509

RESULT 37
US-08-974-549A-219

Sequence 219, Application US/08974549A
Patent No. 6166178

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin B.

APPLICANT: Andrews, William H.

TITLE OF INVENTION: Human Telomerase Catalytic Subunit

NUMBER OF SEQUENCES: 727

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/974,549A

FILING DATE: 19-NOV-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/911,312

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/912,951

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/915,503

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17618

FILING DATE: 01-OCT-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17865

FILING DATE: 01-OCT-1997

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph Ted

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 219:
SEQUENCE CHARACTERISTICS:
LENGTH: 719 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-974-549A-219

Query Match 10.3%; Score 6; DB 4; Length 719;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 DEIKGH 18
|||||
DB 504 DEIKGH 509

RESULT 38
US-08-854-050-7
Sequence 7, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morlin, Gregg B.
APPLICANT: Harley, Calvin
TITLE OF INVENTION: No. 6261836el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
FILING DATE: 01-OCT-1996
APPLICATION NUMBER: US 08/724,643
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 719 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-854-050-7

Query Match 10.3%; Score 6; DB 4; Length 719;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 DEIKGH 18
|||||
DB 504 DEIKGH 509

RESULT 39
US-09-430-323-7
Sequence 7, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morlin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-OCT-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 719 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-430-323-7

Query Match 10.3%; Score 6; DB 4; Length 719;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 DEIKGH 18
|||||
DB 504 DEIKGH 509

RESULT 40
US-08-472-659-7
Sequence 7, Application US/08472659
Patent No. 5831030
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURODOKA, No. 5831030uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5831030unhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5831030oml
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANER
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-472-659-7

Query Match 8.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29 LYDAK 33
|||||
DB 1 LYDAK 5

RESULT 41
US-08-474-661-7
Sequence 7, Application US/08474661
Patent No. 5874253
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURODOKA, No. 5874253uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5874253unhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5874253oml
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,661
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANER
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-474-661-7

Query Match 8.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29 LYDAK 33
|||||
DB 1 LYDAK 5

RESULT 42
US-08-611-977-7

Sequence 7, Application US/08611977
Patent No. 5972886
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUBOBUKA, No. 5972886uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5972886uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAUCHI, Kozo
APPLICANT: YAMAGUCHI, No. 5972886omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,977
FILING DATE: 06-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-611-977-7

Query Match 8.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 LYDAK 33
11111
DB 1 LYDAK 5

RESULT 43
US-08-340-283-146
Sequence 146, Application US/08340283
Patent No. 5861318
GENERAL INFORMATION:
APPLICANT: Elhammer, Ake P.
TITLE OF INVENTION: A SCINTILLATION PROXIMITY ASSAY FOR
TITLE OF INVENTION: N-ACETYLGLACTOSAMINYLTRANSFERASE ACTIVITY
NUMBER OF SEQUENCES: 205

CORRESPONDENCE ADDRESS:
ADDRESSEE: Pharmacia and Upjohn, Inc., Intellect. Prop. Law
ADDRESS: (1920-32-1)
STREET: 301 Henrietta Street
CITY: Kalamazoo
STATE: Michigan
COUNTRY: U.S.A.
ZIP: 49001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,283
FILING DATE:
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Woolton, Thomas A.
REGISTRATION NUMBER: 35,004
REFERENCE/DOCKET NUMBER: 4828
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 385-7914
TELEFAX: (616) 385-6897
TEXT: 224401
INFORMATION FOR SEQ ID NO: 146:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-340-283-146

Query Match 8.6%; Score 5; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 PHOIS 24
11111
DB 1 PHOIS 5

RESULT 44
US-08-383-753-24
Sequence 24, Application US/08383753
Patent No. 5723584
GENERAL INFORMATION:
APPLICANT: Schatz, Peter J.
TITLE OF INVENTION: Biotinylation of Proteins
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourile and Crew
STREET: One Market Plaza, Stewart Tower
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/383,753
FILING DATE: 03-FEB-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,991
FILING DATE: 30-JUL-1993

Db 11111
10 KTVLL 14

Search completed: November 5, 2002, 05:28:46
Job time : 27 secs

ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 1038.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-586-772-24

Query Match 8.6%; Score 5; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 KTVLL 10
11111
Db 10 KTVLL 14

RESULT 45
US-08-586-772-24
Sequence 24, Application US/08586772
Patent No. 3874239
GENERAL INFORMATION:
APPLICANT: Schatz, Peter J.
TITLE OF INVENTION: Biotinylation of Proteins
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Hourie and Crew
STREET: One Market Plaza, Stewart Tower
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/586,772
FILING DATE: 03-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,991
FILING DATE: 30-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 1038.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-586-772-24

Query Match 8.6%; Score 5; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 KTVLL 10

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OK protein - nucleic search, using frame_plus_p2n model

Run on: November 5, 2002, 03:56:25 ; Search time 64 Seconds
(without alignments)
222.605 Million cell updates/sec

Title: US-09-727-892A-99

Perfect score: 58

Sequence: 1 MERRKYVLLYCEIKGHFP.....YERNLFTKKYATIEYIKEL 58

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 383533 seqs, 122816752 residues

Word size: 1

Total number of hits satisfying chosen parameters: 763576

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgn2_1/USPRO.spool/US09727892/rnat_01112002_185928_4878/app.query.fasta.1.199
-DB=issued.patents_NA -QPM=fastap -SUFFIX=NAOLIG.rni -MINMATCH=0.1 -LOOPI=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR.SCORE=quality -THR.MIN=1 -ALIGN=45 -MODE=LOCAL
-USER=US09727892.cgn2_1.1.13 -runat_01112002_185928_4878 -NCPU=6 -ICPU=3
-NO_XLPRY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -YGAPOP=60 -Fgapop=6 -Fgapext=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents.NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
C 1	7	12.1	3186 1	US-07-688-352C-23 Sequence 23, Appl
C 2	7	12.1	3186 2	US-08-474-379C-23 Sequence 23, Appl
C 3	7	12.1	3186 3	US-09-146-249A-23 Sequence 23, Appl
C 4	7	12.1	3186 4	US-08-206-188B-23 Sequence 23, Appl
C 5	7	12.1	3186 5	PCT-US91-02714-22 Sequence 28, Appl
C 6	7	12.1	3585 1	US-08-680-326-28 Sequence 1, Appl
C 7	7	12.1	26700 1	US-08-472-217-1 Sequence 5, Appl
C 8	7	12.1	26700 2	US-08-488-199-5 Sequence 1, Appl
C 9	7	12.1	26700 3	US-08-760-534A-1 Sequence 1, Appl
C 10	6	10.3	33 1	US-08-138-608-33 Sequence 33, Appl
C 11	6	10.3	33 4	US-09-232-479-30 Sequence 30, Appl
C 12	6	10.3	690 3	US-09-064-703-6 Sequence 6, Appl

C	13	6	10.3	988 4	US-08-642-274D-53	Sequence 53, Appl
C 14	6 <td>10.3<td>988 4<td>US-08-952-014C-53<td>Sequence 53, Appl</td></td></td></td>	10.3 <td>988 4<td>US-08-952-014C-53<td>Sequence 53, Appl</td></td></td>	988 4 <td>US-08-952-014C-53<td>Sequence 53, Appl</td></td>	US-08-952-014C-53 <td>Sequence 53, Appl</td>	Sequence 53, Appl	
C 15	6 <td>10.3<td>1092 4<td>US-09-227-357-35<td>Sequence 35, Appl</td></td></td></td>	10.3 <td>1092 4<td>US-09-227-357-35<td>Sequence 35, Appl</td></td></td>	1092 4 <td>US-09-227-357-35<td>Sequence 35, Appl</td></td>	US-09-227-357-35 <td>Sequence 35, Appl</td>	Sequence 35, Appl	
C 16	6 <td>10.3<td>1260 1<td>US-08-599-252-79<td>Sequence 79, Appl</td></td></td></td>	10.3 <td>1260 1<td>US-08-599-252-79<td>Sequence 79, Appl</td></td></td>	1260 1 <td>US-08-599-252-79<td>Sequence 79, Appl</td></td>	US-08-599-252-79 <td>Sequence 79, Appl</td>	Sequence 79, Appl	
C 17	6 <td>10.3<td>1260 1<td>US-08-436-074-52<td>Sequence 52, Appl</td></td></td></td>	10.3 <td>1260 1<td>US-08-436-074-52<td>Sequence 52, Appl</td></td></td>	1260 1 <td>US-08-436-074-52<td>Sequence 52, Appl</td></td>	US-08-436-074-52 <td>Sequence 52, Appl</td>	Sequence 52, Appl	
C 18	6 <td>10.3<td>1260 5<td>PCT-US96-06352-79<td>Sequence 79, Appl</td></td></td></td>	10.3 <td>1260 5<td>PCT-US96-06352-79<td>Sequence 79, Appl</td></td></td>	1260 5 <td>PCT-US96-06352-79<td>Sequence 79, Appl</td></td>	PCT-US96-06352-79 <td>Sequence 79, Appl</td>	Sequence 79, Appl	
C 19	6 <td>10.3<td>1260 1<td>PCT-US96-06583-79<td>Sequence 67, Appl</td></td></td></td>	10.3 <td>1260 1<td>PCT-US96-06583-79<td>Sequence 67, Appl</td></td></td>	1260 1 <td>PCT-US96-06583-79<td>Sequence 67, Appl</td></td>	PCT-US96-06583-79 <td>Sequence 67, Appl</td>	Sequence 67, Appl	
C 20	6 <td>10.3<td>1269 1<td>US-08-680-726A-67<td>Sequence 67, Appl</td></td></td></td>	10.3 <td>1269 1<td>US-08-680-726A-67<td>Sequence 67, Appl</td></td></td>	1269 1 <td>US-08-680-726A-67<td>Sequence 67, Appl</td></td>	US-08-680-726A-67 <td>Sequence 67, Appl</td>	Sequence 67, Appl	
C 21	6 <td>10.3<td>1269 3<td>US-09-092-409-67<td>Sequence 67, Appl</td></td></td></td>	10.3 <td>1269 3<td>US-09-092-409-67<td>Sequence 67, Appl</td></td></td>	1269 3 <td>US-09-092-409-67<td>Sequence 67, Appl</td></td>	US-09-092-409-67 <td>Sequence 67, Appl</td>	Sequence 67, Appl	
C 22	6 <td>10.3<td>1351 2<td>US-08-401-068-13<td>Sequence 13, Appl</td></td></td></td>	10.3 <td>1351 2<td>US-08-401-068-13<td>Sequence 13, Appl</td></td></td>	1351 2 <td>US-08-401-068-13<td>Sequence 13, Appl</td></td>	US-08-401-068-13 <td>Sequence 13, Appl</td>	Sequence 13, Appl	
C 23	6 <td>10.3<td>1351 2<td>US-08-846-338-13<td>Sequence 13, Appl</td></td></td></td>	10.3 <td>1351 2<td>US-08-846-338-13<td>Sequence 13, Appl</td></td></td>	1351 2 <td>US-08-846-338-13<td>Sequence 13, Appl</td></td>	US-08-846-338-13 <td>Sequence 13, Appl</td>	Sequence 13, Appl	
C 24	6 <td>10.3<td>1524 1<td>US-08-409-122-1<td>Sequence 1, Appl</td></td></td></td>	10.3 <td>1524 1<td>US-08-409-122-1<td>Sequence 1, Appl</td></td></td>	1524 1 <td>US-08-409-122-1<td>Sequence 1, Appl</td></td>	US-08-409-122-1 <td>Sequence 1, Appl</td>	Sequence 1, Appl	
C 25	6 <td>10.3<td>1524 2<td>US-08-408-669-1<td>Sequence 1, Appl</td></td></td></td>	10.3 <td>1524 2<td>US-08-408-669-1<td>Sequence 1, Appl</td></td></td>	1524 2 <td>US-08-408-669-1<td>Sequence 1, Appl</td></td>	US-08-408-669-1 <td>Sequence 1, Appl</td>	Sequence 1, Appl	
C 26	6 <td>10.3<td>1590 3<td>US-09-064-703-1<td>Sequence 1, Appl</td></td></td></td>	10.3 <td>1590 3<td>US-09-064-703-1<td>Sequence 1, Appl</td></td></td>	1590 3 <td>US-09-064-703-1<td>Sequence 1, Appl</td></td>	US-09-064-703-1 <td>Sequence 1, Appl</td>	Sequence 1, Appl	
C 27	6 <td>10.3<td>1605 3<td>US-09-064-703-5<td>Sequence 5, Appl</td></td></td></td>	10.3 <td>1605 3<td>US-09-064-703-5<td>Sequence 5, Appl</td></td></td>	1605 3 <td>US-09-064-703-5<td>Sequence 5, Appl</td></td>	US-09-064-703-5 <td>Sequence 5, Appl</td>	Sequence 5, Appl	
C 28	6 <td>10.3<td>1664 1<td>US-07-863-169A-6<td>Sequence 6, Appl</td></td></td></td>	10.3 <td>1664 1<td>US-07-863-169A-6<td>Sequence 6, Appl</td></td></td>	1664 1 <td>US-07-863-169A-6<td>Sequence 6, Appl</td></td>	US-07-863-169A-6 <td>Sequence 6, Appl</td>	Sequence 6, Appl	
C 29	6 <td>10.3<td>1664 2<td>US-08-429-964-6<td>Sequence 6, Appl</td></td></td></td>	10.3 <td>1664 2<td>US-08-429-964-6<td>Sequence 6, Appl</td></td></td>	1664 2 <td>US-08-429-964-6<td>Sequence 6, Appl</td></td>	US-08-429-964-6 <td>Sequence 6, Appl</td>	Sequence 6, Appl	
C 30	6 <td>10.3<td>1664 3<td>US-07-935-087-6<td>Sequence 6, Appl</td></td></td></td>	10.3 <td>1664 3<td>US-07-935-087-6<td>Sequence 6, Appl</td></td></td>	1664 3 <td>US-07-935-087-6<td>Sequence 6, Appl</td></td>	US-07-935-087-6 <td>Sequence 6, Appl</td>	Sequence 6, Appl	
C 31	6 <td>10.3<td>1664 5<td>PCT-US93-08062-6<td>Sequence 6, Appl</td></td></td></td>	10.3 <td>1664 5<td>PCT-US93-08062-6<td>Sequence 6, Appl</td></td></td>	1664 5 <td>PCT-US93-08062-6<td>Sequence 6, Appl</td></td>	PCT-US93-08062-6 <td>Sequence 6, Appl</td>	Sequence 6, Appl	
C 32	6 <td>10.3<td>1699 4<td>US-09-058-260-13<td>Sequence 13, Appl</td></td></td></td>	10.3 <td>1699 4<td>US-09-058-260-13<td>Sequence 13, Appl</td></td></td>	1699 4 <td>US-09-058-260-13<td>Sequence 13, Appl</td></td>	US-09-058-260-13 <td>Sequence 13, Appl</td>	Sequence 13, Appl	
C 33	6 <td>10.3<td>1701 3<td>US-09-064-703-4<td>Sequence 4, Appl</td></td></td></td>	10.3 <td>1701 3<td>US-09-064-703-4<td>Sequence 4, Appl</td></td></td>	1701 3 <td>US-09-064-703-4<td>Sequence 4, Appl</td></td>	US-09-064-703-4 <td>Sequence 4, Appl</td>	Sequence 4, Appl	
C 34	6 <td>10.3<td>1733 4<td>US-09-058-260-21<td>Sequence 21, Appl</td></td></td></td>	10.3 <td>1733 4<td>US-09-058-260-21<td>Sequence 21, Appl</td></td></td>	1733 4 <td>US-09-058-260-21<td>Sequence 21, Appl</td></td>	US-09-058-260-21 <td>Sequence 21, Appl</td>	Sequence 21, Appl	
C 35	6 <td>10.3<td>1756 4<td>US-09-058-260-31<td>Sequence 31, Appl</td></td></td></td>	10.3 <td>1756 4<td>US-09-058-260-31<td>Sequence 31, Appl</td></td></td>	1756 4 <td>US-09-058-260-31<td>Sequence 31, Appl</td></td>	US-09-058-260-31 <td>Sequence 31, Appl</td>	Sequence 31, Appl	
C 36	6 <td>10.3<td>1776 4<td>US-09-058-260-33<td>Sequence 23, Appl</td></td></td></td>	10.3 <td>1776 4<td>US-09-058-260-33<td>Sequence 23, Appl</td></td></td>	1776 4 <td>US-09-058-260-33<td>Sequence 23, Appl</td></td>	US-09-058-260-33 <td>Sequence 23, Appl</td>	Sequence 23, Appl	
C 37	6 <td>10.3<td>1794 2<td>US-08-427-497E-5<td>Sequence 5, Appl</td></td></td></td>	10.3 <td>1794 2<td>US-08-427-497E-5<td>Sequence 5, Appl</td></td></td>	1794 2 <td>US-08-427-497E-5<td>Sequence 5, Appl</td></td>	US-08-427-497E-5 <td>Sequence 5, Appl</td>	Sequence 5, Appl	
C 38	6 <td>10.3<td>1895 4<td>US-09-058-260-3<td>Sequence 3, Appl</td></td></td></td>	10.3 <td>1895 4<td>US-09-058-260-3<td>Sequence 3, Appl</td></td></td>	1895 4 <td>US-09-058-260-3<td>Sequence 3, Appl</td></td>	US-09-058-260-3 <td>Sequence 3, Appl</td>	Sequence 3, Appl	
C 39	6 <td>10.3<td>1925 4<td>US-09-058-260-17<td>Sequence 17, Appl</td></td></td></td>	10.3 <td>1925 4<td>US-09-058-260-17<td>Sequence 17, Appl</td></td></td>	1925 4 <td>US-09-058-260-17<td>Sequence 17, Appl</td></td>	US-09-058-260-17 <td>Sequence 17, Appl</td>	Sequence 17, Appl	
C 40	6 <td>10.3<td>1950 2<td>US-08-472-659-30<td>Sequence 30, Appl</td></td></td></td>	10.3 <td>1950 2<td>US-08-472-659-30<td>Sequence 30, Appl</td></td></td>	1950 2 <td>US-08-472-659-30<td>Sequence 30, Appl</td></td>	US-08-472-659-30 <td>Sequence 30, Appl</td>	Sequence 30, Appl	
C 41	6 <td>10.3<td>1950 2<td>US-08-474-661-30<td>Sequence 30, Appl</td></td></td></td>	10.3 <td>1950 2<td>US-08-474-661-30<td>Sequence 30, Appl</td></td></td>	1950 2 <td>US-08-474-661-30<td>Sequence 30, Appl</td></td>	US-08-474-661-30 <td>Sequence 30, Appl</td>	Sequence 30, Appl	
C 42	6 <td>10.3<td>1950 2<td>US-08-611-977-30<td>Sequence 30, Appl</td></td></td></td>	10.3 <td>1950 2<td>US-08-611-977-30<td>Sequence 30, Appl</td></td></td>	1950 2 <td>US-08-611-977-30<td>Sequence 30, Appl</td></td>	US-08-611-977-30 <td>Sequence 30, Appl</td>	Sequence 30, Appl	
C 43	6 <td>10.3<td>1952 4<td>US-09-058-260-5<td>Sequence 5, Appl</td></td></td></td>	10.3 <td>1952 4<td>US-09-058-260-5<td>Sequence 5, Appl</td></td></td>	1952 4 <td>US-09-058-260-5<td>Sequence 5, Appl</td></td>	US-09-058-260-5 <td>Sequence 5, Appl</td>	Sequence 5, Appl	
C 44	6 <td>10.3<td>1957 4<td>US-09-058-260-19<td>Sequence 19, Appl</td></td></td></td>	10.3 <td>1957 4<td>US-09-058-260-19<td>Sequence 19, Appl</td></td></td>	1957 4 <td>US-09-058-260-19<td>Sequence 19, Appl</td></td>	US-09-058-260-19 <td>Sequence 19, Appl</td>	Sequence 19, Appl	
C 45	6 <td>10.3<td>1996 3<td>US-08-946-026-15<td>Sequence 15, Appl</td></td></td></td>	10.3 <td>1996 3<td>US-08-946-026-15<td>Sequence 15, Appl</td></td></td>	1996 3 <td>US-08-946-026-15<td>Sequence 15, Appl</td></td>	US-08-946-026-15 <td>Sequence 15, Appl</td>	Sequence 15, Appl	

ALIGNMENTS

RESULT 1
US-07-688-352C-23/C
Sequence 23, Application US/07688352C
Patent No. 5527896
GENERAL INFORMATION:
APPLICANT: Colicelli, John J.
TITLE OF INVENTION: Cloning by Complementation and Related
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Bicknell
STREET: Two First National Plaza, 20 South Clark
STREET: Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/688,352C
FILING DATE: 19910419
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,715
FILING DATE: 20-Apr-1990
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25447
REFERENCE/DOCKET NUMBER: 27805/30197
TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 3186 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 139..2348
US-07-688-352C-23

Alignment Scores:
Pred. No.: 92.6
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 12.07%
DB: 1
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0
Length: 3186

US-09-727-892a-99 (1-58) x US-07-688-352C-23 (1-3186)

OY 16 LysGlyHisPheProHisGln 22
|||||
DB 442 AAGGCCATTTCACATCA 422

RESULT 2
US-08-474-379C-23/C
Sequence 23, Application US/08474379C
Patent No. 5977305
GENERAL INFORMATION:
APPLICANT: Wiegler, Michael H.
APPLICANT: Colicelli, John J.
TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED
TITLE OF INVENTION: PROCESSES
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,379C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,715
FILING DATE: 20-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/206,188
FILING DATE: 01-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/688,352
FILING DATE: 19-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 27866/32771
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:

LENGTH: 3186 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 139..2346
US-08-474-379C-23

Alignment Scores:
Pred. No.: 92.6
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 12.07%
DB: 2
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0
Length: 3186

US-09-727-892a-99 (1-58) x US-08-474-379C-23 (1-3186)

OY 16 LysGlyHisPheProHisGln 22
|||||
DB 442 AAGGCCATTTCACATCA 422

RESULT 3
US-09-146-249A-23/C
Sequence 23, Application US/09146249A
Patent No. 6069240
GENERAL INFORMATION:
APPLICANT: Wiegler, Michael H.
APPLICANT: Colicelli, John J.
TITLE OF INVENTION: Cloning by Complementation and Related
TITLE OF INVENTION: Processes
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/146,249A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,715
FILING DATE: 20-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 3186 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 139..2348
US-09-146-249A-23

Alignment Scores:

Pred. No.: 92.6 Length: 3186
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 DB: 3 Gaps: 0

US-09-727-892a-99 (1-58) x US-09-146-249a-23 (1-3186)

OY 16 LysGlyHisPheProHisGln 22

Db 442 AAGGCCATTTCACATCAA 422

RESULT 4

US-08-206-188B-23/C
 ; Sequence 23, Application US/08206188B
 ; Patent No. 6100025

GENERAL INFORMATION:

APPLICANT: Wigler, Michael H.

APPLICANT: Colicelli, John J.

TITLE OF INVENTION: Cloning by Complementation and Related Processes

NUMBER OF SEQUENCES: 84

CORRESPONDENCE ADDRESS:

ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/206,188B

FILING DATE: 01-MAR-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/511,715

FILING DATE: 20-APR-1990

ATTORNEY/AGENT INFORMATION:

NAME: Clough, David W.

REGISTRATION NUMBER: 36107

TELEPHONE: 312/474-6300

TELEFAX: 312-474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 3186 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 139..2348

US-08-206-188B-23

Alignment Scores:

Pred. No.: 92.6 Length: 3186

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 12.07% Indels: 0

DB: 3 Gaps: 0

US-09-727-892a-99 (1-58) x US-08-206-188B-23 (1-3186)

OY 16 LysGlyHisPheProHisGln 22

|||||

Db 442 AAGGCCATTTCACATCAA 422

RESULT 5

PCT-US91-02714-22/C
 ; Sequence 22, Application PC/TUS9102714

GENERAL INFORMATION:

APPLICANT: Wigler, Michael H.

APPLICANT: Colicelli, John J.

TITLE OF INVENTION: Cloning by Complementation and Related Processes

NUMBER OF SEQUENCES: 55

CORRESPONDENCE ADDRESS:

ADDRESS: Marshall, O'Toole, Gerstein, Murray &

STREET: Two First National Plaza, 20 South Clark

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/02714

FILING DATE: 19910419

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/511,715

FILING DATE: 20-APR-1990

ATTORNEY/AGENT INFORMATION:

NAME: Borun, Michael F.

REGISTRATION NUMBER: 25447

REFERENCE/DOCKET NUMBER: 27805/30197

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 346-5750

TELEFAX: (312) 984-9740

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 3186 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 139..2348

PCT-US91-02714-22

Alignment Scores:

Pred. No.: 92.6 Length: 3186

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 12.07% Indels: 0

DB: 5 Gaps: 0

US-09-727-892a-99 (1-58) x PCT-US91-02714-22 (1-3186)

OY 16 LysGlyHisPheProHisGln 22

Db 442 AAGGCCATTTCACATCAA 422

RESULT 6

US-08-680-326-28/C

; Sequence 28, Application US/08680326

; Patent No. 5925733

; GENERAL INFORMATION:

; APPLICANT: ROSE, TIMOTHY M.

; APPLICANT: BOSCH, MARNIX

APPLICANT: STRAND, KURT
APPLICANT: TODARO, GEORGE J.
TITLE OF INVENTION: DNA POLYMERASE OF GAMMA HERPES VIRUSES
TITLE OF INVENTION: ASSOCIATED WITH KAPOSI'S SARCOMA AND RETROPERITONEAL
TITLE OF INVENTION: FIBROMATOSIS
NUMBER OF SEQUENCES: 152
CORRESPONDENCE ADDRESS:
ADDRESS: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/680.326
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Schiff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20001.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 3585 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-680-326-28

Alignment Scores:
Pred. No.: 103
Score: 7.00
Length: 3585
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 12.07%
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-727-892a-99 (1-58) x US-08-680-326-28 (1-3585)

OY 37 SerTyTyGluTyAsnLeu 43
|||||

Db 2458 AGTATATGATATGATATCTC 2438

RESULT 7
US-08-472-217-1/c
Sequence 1, Application US/08472217
Patent No. 5726058
GENERAL INFORMATION:
APPLICANT: Alanen-Kurki, Leena
APPLICANT: Auvainen, Petri
APPLICANT: Jaakkola, Panu
APPLICANT: Jalkanen, Markku
APPLICANT: Lepp, Sirpa
APPLICANT: Maki, Markku
APPLICANT: Viikari, Tapani
TITLE OF INVENTION: Syndecan Stimulation Of Cellular
TITLE OF INVENTION: Differentiation
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.

COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472.217
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/206,186
FILING DATE: 07-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/988,427
FILING DATE: 01-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Cimada, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 1102.0050003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 26700 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: join(4378..4443, 22026..22106, 23001..23483,
LOCATION: 23905..24039, 24251..24418)
US-08-472-217-1

Alignment Scores:
Pred. No.: 654
Score: 7.00
Length: 26700
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 12.07%
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-727-892a-99 (1-58) x US-08-472-217-1 (1-26700)

OY 45 ThrLysLysTyraLysLeu 51
|||||

Db 17464 ACAAATAATGCGTACATA 17444

RESULT 8
US-08-488-199-5/c
Sequence 5, Application US/08488199
Patent No. 5851993
GENERAL INFORMATION:
APPLICANT: Jalkanen, Markku
APPLICANT: Maki, Markku
TITLE OF INVENTION: Suppression of Tumor Cell Growth By
TITLE OF INVENTION: Syndecan-1 Ectodomain
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
STREET: 1100 New York Ave., NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,199
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/258,862
FILING DATE: 13-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Cimbal, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 1102, 0130001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 26700 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 4378..4443
FEATURE:
NAME/KEY: CDS
LOCATION: 22026..22107
FEATURE:
NAME/KEY: CDS^b
LOCATION: 23002..23483
FEATURE:
NAME/KEY: CDS
LOCATION: 23905..24040
FEATURE:
NAME/KEY: CDS
LOCATION: 24252..24418
US-08-488-199-5

Alignment Scores:
Pred. No.: 654
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 12.07%
DB: 2
Gaps: 0

US-09-727-892a-99 (1-58) x US-08-488-199-5 (1-26700)

OY 45 ThrlyslsTYrAlaTYrIle 51
|||||
Db 17464 ACAAAAATATGCGTACATA 17444

RESULT 9
US-08-760-534A-1/c
Sequence 1, Application US/08760534A
Patent No. 6017727
GENERAL INFORMATION:
APPLICANT: JALKANEN, MARKKU
APPLICANT: JAAKKOLA, PANU
APPLICANT: VIHINEN, TAPANI
TITLE OF INVENTION: SYNDSCAN ENHANCER ELEMENT AND SYNDSCAN
TITLE OF INVENTION: SYNDSCAN ENHANCER ELEMENT AND SYNDSCAN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: US
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/760,534A
FILING DATE: 02-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/206,186
FILING DATE: 07-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FI93/00514
FILING DATE: 01-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: CIMBALA, MICHELE A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 1708, 0050004/MAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 26700 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(4378..4443, 22026..22106, 23001..23483,
LOCATION: 23905..24039, 24251..24418)
US-08-760-534A-1

Alignment Scores:
Pred. No.: 654
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 12.07%
DB: 3
Gaps: 0

US-09-727-892a-99 (1-58) x US-08-760-534A-1 (1-26700)

OY 45 ThrlyslsTYrAlaTYrIle 51
|||||
Db 17464 ACAAAAATATGCGTACATA 17444

RESULT 10
US-08-138-608-33/c
Sequence 33, Application US/08138608
Patent No. 5407795
GENERAL INFORMATION:
APPLICANT: Kolberg, Janice A.
APPLICANT: Shen, Lu-Ping
APPLICANT: Urdea, Michael S.
TITLE OF INVENTION: CMV PROBES FOR USE IN SOLUTION
TITLE OF INVENTION: PHASE SANDWICH HYBRIDIZATION ASSAYS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/138,608
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/813,590
FILING DATE: 23-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. Clotelli
REGISTRATION NUMBER: 21,013
REFERENCE/DOCKET NUMBER: 22300-20236.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-813-5600
TELEFAX: 415-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-138-608-33

Alignment Scores:
Pred. No.: 16.6 Length: 33
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 1 Gaps: 0

US-09-727-892a-99 (1-58) x US-08-138-608-33 (1-33)

QY 40 GluTYRAsnLeuPheThr 45
Db 33 GAGTACACACTGTTACG 16

RESULT 11
US-09-232-479-30
Sequence 30, Application US/09232479
Patent No. 6221362
GENERAL INFORMATION:
APPLICANT: AUDONNET, JEAN-CHRISTOPHE
APPLICANT: BOUCHARDON, ANNABELLE
APPLICANT: RIVIERE, MICHEL
TITLE OF INVENTION: AVIAN POLYNUCLEOTIDE VACCINE FORMULA
FILE REFERENCE: 454313-2260
CURRENT FILING DATE: 1999-01-15
EARLIER APPLICATION NUMBER: US/09/232,479
EARLIER FILING DATE: 1996-07-19
EARLIER APPLICATION NUMBER: 96/09339
EARLIER FILING DATE: 1997-07-16
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 30
LENGTH: 33
TYPE: DNA
ORGANISM: chicken infectious laryngotracheitis virus
US-09-232-479-30

Alignment Scores:
Pred. No.: 16.6 Length: 33
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 4 Gaps: 0

US-09-727-892a-99 (1-58) x US-09-232-479-30 (1-33)

QY 27 GluAspLeuTYRAspAla 32
Db 2 GAAAGATCTTACGATGCT 19

RESULT 12

US-09-064-703-6
Sequence 6, Application US/09064703
Patent No. 6033894
GENERAL INFORMATION:
APPLICANT: Craik, Charles S.
APPLICANT: Unal, Ayce
TITLE OF INVENTION: Ganem, Donald E.
TITLE OF INVENTION: Kaposi's Syndrome Herpesvirus
TITLE OF INVENTION: Protease and Assembly Protein Compositions and
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/064,703
FILING DATE: 22-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/044,152
FILING DATE: 22-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Petithory, Joanne R
REGISTRATION NUMBER: P42,995
REFERENCE/DOCKET NUMBER: 2002-0002.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0980
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 690 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: KSHV pr coding sequence
US-09-064-703-6

Alignment Scores:
Pred. No.: 272 Length: 690
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 3 Gaps: 0

US-09-727-892a-99 (1-58) x US-09-064-703-6 (1-690)

QY 19 PheProHisGlnIleSer 24
Db 551 TTTCACACCAATTTG 568

RESULT 13
US-08-642-274D-53/C
Sequence 53, Application US/08642274D
Patent No. 6200749
GENERAL INFORMATION:
APPLICANT: Shiloh, Yosef
TITLE OF INVENTION: MOTATED FORMS OF THE ATAXIA-TELANGIECTASIA GENE AND METHOD TO
FILE REFERENCE: 229000033
CURRENT APPLICATION NUMBER: US/08/642,274D
CURRENT FILING DATE: 1996-05-03
NUMBER OF SEQ ID NOS: 220
SOFTWARE: Patentin Ver. 2.1

```
; SEQ ID NO 53
; LENGTH: 988
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: genomic
US-08-642-274D-53

Alignment Scores:
Pred. No.: 378 Length: 988
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 4 Gaps: 0

US-09-727-892a-99 (1-58) x US-08-642-274D-53 (1-988)

OY 6 LysThrValleuLeuTyr 11
    |||||
DB 156 AAGACAGTGCCTCTCTAT 139

RESULT 14
US-08-952-014C-53/C
; Sequence 53, Application US/08952014C
; Patent No. 6265158
; GENERAL INFORMATION:
; APPLICANT: Shiloh, Yosef
; TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE AND ITS
; TITLE OF INVENTION: GENOMIC ORGANIZATION
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 6265158thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/952,014C
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,995
; REFERENCE/DOCKET NUMBER: 2290,00028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 810-539-5050
; TELEFAX: 810-539-5055
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 988 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-952-014C-53

Alignment Scores:
Pred. No.: 378 Length: 988
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 4 Gaps: 0

US-09-727-892a-99 (1-58) x US-08-952-014C-53 (1-988)
```

```
OY 6 LysThrValleuLeuTyr 11
    |||||
DB 156 AAGACAGTGCCTCTCTAT 139

RESULT 15
US-09-227-357-35
; Sequence 35, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
```

EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 35
LENGTH: 1092
TYPE: DNA
ORGANISM: Homo sapiens
US-09-727-357-35

Alignment Scores:
Pred. No.: 415 Length: 1092
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 4 Gaps: 0

US-09-727-892a-99 (1-58) x US-09-227-357-35 (1-1092)

OY 16 LysGlyHisPheProHis 21
Db 76 AAAGACATTTCCCCAC 93

RESULT 16
US-08-599-252-79
Sequence 79, Application US/08599252
Patent No. 5705343
GENERAL INFORMATION:
APPLICANT: DRAVNA, DENNIS T.
APPLICANT: FEDER, JOHN N.
APPLICANT: GINRKE, ANDREAS
APPLICANT: KIMMEL, BRUCE E.
APPLICANT: THOMAS, WINSTON J.
APPLICANT: WOLFF, ROGER K.
TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
TITLE OF INVENTION: HEMOCHROMATOSIS
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/599,252
FILING DATE: 09-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 9053-0001.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 1260 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-599-252-79

Alignment Scores:
Pred. No.: 473 Length: 1260
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 1 Gaps: 0

US-09-727-892a-99 (1-58) x US-08-599-252-79 (1-1260)

OY 39 TyrGluTyrAsnLeuphe 44
Db 96 TATGATATCAATTTATTT 113

RESULT 17
US-08-436-074-52
Sequence 52, Application US/08436074
Patent No. 5753438
GENERAL INFORMATION:
APPLICANT: DRAVNA, DENNIS T.
APPLICANT: FEDER, JOHN N.
APPLICANT: GINRKE, ANDREAS
APPLICANT: KIMMEL, BRUCE E.
APPLICANT: THOMAS, WINSTON J.
APPLICANT: WOLFF, ROGER K.
TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
TITLE OF INVENTION: HEMOCHROMATOSIS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,074
FILING DATE: 08-MAY-1995
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 9053-0001.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 1260 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-436-074-52

Alignment Scores:
Pred. No.: 473 Length: 1260
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 1 Gaps: 0

US-09-727-892a-99 (1-58) x US-08-436-074-52 (1-1260)

OY 39 TyrGluTyrAsnLeuphe 44


```

Db          96 TATGAATACAAATTATT 113
|||||
RESULT 18
PCT-US96-06352-79
; Sequence 79, Application PC/TUS9606352
; GENERAL INFORMATION:
APPLICANT: DRAVNA, DENNIS T.
APPLICANT: FEDER, JOHN N.
APPLICANT: GINTKE, ANDREAS
APPLICANT: KIMMEL, BROCE E.
APPLICANT: THOMAS, WINGSTON J.
APPLICANT: WOLFF, ROGER K.
TITLE OF INVENTION: METHOD TO DIAGNOSE HEMEDITARY
TITLE OF INVENTION: HEROCROMATOSIS
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESSES:
ADDRESSER: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
City: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/06352
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,252
FILING DATE: 09-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 9053-0001.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 1260 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US96-06352-79

Alignment Scores:
Pred. No.:      473      Length:    1260
Score:           6.00     Matches:     6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches:   0
Query Match:       10.34% Indels:        0
DB:                5      Gaps:         0

US-09-727-892A-99 (1-58) x PCT-US96-06352-79 (1-1260)
OY          39 TytGluTyrAsnleuphe 44
|||||
Db          96 TATGAATACAAATTATT 113

RESULT 19
PCT-US96-06583-79
; Sequence 79, Application PC/TUS9606583
; GENERAL INFORMATION:
APPLICANT: DRAVNA, DENNIS T.
APPLICANT: FEDER, JOHN N.
APPLICANT: GINTKE, ANDREAS
APPLICANT: KIMMEL, BROCE E.

```

```

APPLICANT: THOMAS, WINSTON J.
APPLICANT: WOLFE, ROGER K.
TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
TITLE OF INVENTION: HEMOCHROMATOSIS
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
City: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT-US96/06583
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,252
FILING DATE: 09-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: MORASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 9053-0001.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 1260 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US96-06583-79

Alignment Scores:
Pred. No.:      473          Length:    1260
Score:           6.00         Matches:     6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:     10.34%       Indels:      0
DB:              5           Gaps:        0

US-09-727-892A-99 (1-58) x PCT-US96-06583-79 (1-1260)
QY      39   TyGtUgTyAsnuIeuphe 44
          |||||||
Db      96   TATGATAACATTATTT 113

RESULT 20
US-08-680-726A-67
Sequence 67, Application US/08680726A
Patent No. 5804197
GENERAL INFORMATION:
APPLICANT: Haanes, Elizabeth J.
APPLICANT: Frank, Rexann S.
TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUSES
NUMBER OF SEQUENCES: 92
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
City: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/680,726A
FILING DATE: 12-JUL-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-46-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 1269 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1266
US-08-680-726A-67

Alignment Scores:
Pred. No.: 476 Length: 1269
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
Gaps: 0
DB: 1

US-09-727-892A-99 (1-58) x US-08-680-726A-67 (1-1269)
QY 40 GltYrAsnLeuphThr 45
Db 1192 GAATATATTATTATTACA 1209

RESULT 21
US-09-092-409-67
Sequence 67, Application US/09092409
Patent No. 6139478
GENERAL INFORMATION:
APPLICANT: Haanes, Elizabeth J.
APPLICANT: Frank, Rexann S.
TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUSES
NUMBER OF SEQUENCES: 92
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/092,409
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/680,726
FILING DATE: 12-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-46-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700

TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 1269 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1266
US-09-092-409-67

Alignment Scores:
Pred. No.: 476 Length: 1269
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
Gaps: 0
DB: 3

US-09-727-892A-99 (1-58) x US-09-092-409-67 (1-1269)
QY 40 GltYrAsnLeuphThr 45
Db 1192 GAATATATTATTATTACA 1209

RESULT 22
US-08-401-068-13
Sequence 13, Application US/08401068
Patent No. 5859335
GENERAL INFORMATION:
APPLICANT: Patton, David
TITLE OF INVENTION: Enhanced Biotin Biosynthesis In Plant Tissue
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/401,068
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/351,970
FILING DATE: 08-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1351 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 59..1192
OTHER INFORMATION: /product- "Arabidopsis Biot enzyme"
US-08-401-068-13

Alignment Scores:

Pred. No.: 505 Length: 1351
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 2 Gaps: 0

US-09-727-892a-99 (1-58) x US-08-401-068-13 (1-1351)

QY 53 GIUTYRIIElysgluile 58

DB 530 GAATACATCAAGAATA 547

RESULT 23

US-08-846-338-13

; Sequence 13, Application US/08846338

; Patent No. 5869719

; GENERAL INFORMATION:

; APPLICANT: Patton, David

; TITLE OF INVENTION: Transgenic Plants Having Increased Biotin Content

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 5869719artis Corporation

; STREET: 520 White Plains Road, P.O. Box 2005

; CITY: Tarrytown

; STATE: NY

; COUNTRY: USA

; ZIP: 10591

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: PC-DOS/MS-DOS

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/846,338

; FILING DATE:

; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:

; NAME: Meigs, J. Timothy

; REGISTRATION NUMBER: 38,241

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 919-541-8587

; TELEFAX: 919-541-8689

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1351 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: CDNA

; HYPOTHETICAL: NO

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 59..1192

; OTHER INFORMATION: /product= "Arabidopsis biotin synthase enzyme"

; US-08-846-338-13

Alignment Scores:

Pred. No.: 505 Length: 1351
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 2 Gaps: 0

US-09-727-892a-99 (1-58) x US-08-846-338-13 (1-1351)

QY 53 GIUTYRIIElysgluile 58

DB 530 GAATACATCAAGAATA 547

RESULT 24

US-08-409-122-1/C

; Sequence 1, Application US/08409122

; Patent No. 5820870

; GENERAL INFORMATION:

; APPLICANT: JOYCE, JAMES G.

; APPLICANT: GEORGE, HUGH A.

; APPLICANT: HOFMANN, KATHRYN J.

; APPLICANT: JANSSEN, KATHRYN U.

; APPLICANT: NEPPER, MICHAEL P.

; TITLE OF INVENTION: DNA ENCODING HUMAN PAPILLOMAVIRUS TYPE 18 VACCINE

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CHRISTINE E. CARTY - MERCK & CO., INC.

; STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000

; CITY: RAHWAY

; STATE: NJ

; COUNTRY: US

; ZIP: 07065-0907

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/409,122

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/408,669

; FILING DATE: 22-MAR-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: CARTY, CHRISTINE E

; REGISTRATION NUMBER: 36,099

; REFERENCE/DOCKET NUMBER: 19425

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 908-594-6734

; TELEFAX: 908-594-4720

; TEXAS:

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1524 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: CDNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE:

; ORIGINAL SOURCE:

; US-08-409-122-1

Alignment Scores:

Pred. No.: 564 Length: 1524
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 1 Gaps: 0

US-09-727-892a-99 (1-58) x US-08-409-122-1 (1-1524)

QY 36 TYTSETYTYTgltutyr 41

DB 1189 TATTCATCAATGAAATAT 1172

RESULT 25

US-08-408-669-1/C

; Sequence 1, Application US/08408669

; Patent No. 5840306

; GENERAL INFORMATION:

; APPLICANT: HOFMANN, KATHRYN J.

; APPLICANT: JANSSEN, KATHRYN U.

APPLICANT: NEPPER, MICHAEL P.
TITLE OF INVENTION: DNA ENCODING HUMAN PAPILLOMAVIRUS TYPE 18
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHRISTINE E. CARRY - MERCK & CO., INC.
STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
CITY: RAHWAY
STATE: NJ
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,669
FILING DATE: 22-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: CARRY, CHRISTINE E.
REGISTRATION NUMBER: 36,099
REFERENCE/DOCKET NUMBER: 19424
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-6734
TELEFAX: 908-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1524 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-408-669-1

Alignment Scores:
Pred. No.: 564 Length: 1524
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 2 Gaps: 0

US-09-727-892A-99 (1-58) x US-08-408-669-1 (1-1524)

QY 36 TyTSeTyTyrGluTyr 41
Db 1189 TATTCATACATGATAT 1172

RESULT 26
US-09-064-703-1
Sequence 1, Application US/09064703
Patent No. 6033894
GENERAL INFORMATION:
APPLICANT: Craik, Charles S.
APPLICANT: Unal, Ayce
APPLICANT: Ganem, Donald E.
TITLE OF INVENTION: Kaposi's Syndrome Herpesvirus
TITLE OF INVENTION: Protease and Assembly Protein Compositions and
METHODS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto

STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/064,703
FILING DATE: 22-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/044,152
FILING DATE: 22-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Pettibory, Joanne R.
REGISTRATION NUMBER: P42,995
REFERENCE/DOCKET NUMBER: 2002-0002.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1590 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: KSHV Pr/AP polypotein coding sequence
US-09-064-703-1

Alignment Scores:
Pred. No.: 586 Length: 1590
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 3 Gaps: 0

US-09-727-892A-99 (1-58) x US-09-064-703-1 (1-1590)

QY 19 PheProHisGlnIleSer 24
Db 551 TTTCCACACCAATTTCG 568

RESULT 27
US-09-064-703-5
Sequence 5, Application US/09064703
Patent No. 6033894
GENERAL INFORMATION:
APPLICANT: Craik, Charles S.
APPLICANT: Unal, Ayce
APPLICANT: Ganem, Donald E.
TITLE OF INVENTION: Kaposi's Syndrome Herpesvirus
TITLE OF INVENTION: Protease and Assembly Protein Compositions and
METHODS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/064,703
FILING DATE: 22-APR-1998
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/044,152
FILING DATE: 22-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Peltihory, Joanne R
REGISTRATION NUMBER: P42,995
REFERENCE/DOCKET NUMBER: 2002-0002.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0960
TELEFAX: 650-324-0880
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1605 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: KSHV Pr/AP polyprotein including SEQ ID NO:3
US-09-064-703-5

Alignment Scores:
Pred. No.: 591 Length: 1605
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 3 Gaps: 0

US-09-727-892A-99 (1-58) x US-09-064-703-5 (1-1605)

OY 19 PheProHisGlnIleSer 24
Db 551 TTTCACACCAATTTCG 568

RESULT 28
US-07-863-169A-6/c
Sequence 6, Application US/07863169A
Patent No. 5420245
GENERAL INFORMATION:
APPLICANT: Brown, Michael S.
APPLICANT: Goldstein, Joseph L.
APPLICANT: Reiss, Yuval
TITLE OF INVENTION: Tetrapeptide-Based Inhibitors of Farnesyl
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/863,169A
FILING DATE: 03-APR-1992
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/822,011
FILING DATE: 19-JAN-1992
CLASSIFICATION: 530
APPLICATION NUMBER: US 07/937,893
FILING DATE: 18-APR-1991
CLASSIFICATION: 530
APPLICATION NUMBER: US 615,715
FILING DATE: 20-NOV-1990
CLASSIFICATION: 530
APPLICATION NUMBER: US 510,706
FILING DATE: 18-APR-1990
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UTSD:297/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1664 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-07-863-169A-6

Alignment Scores:
Pred. No.: 611 Length: 1664
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 1 Gaps: 0

US-09-727-892A-99 (1-58) x US-07-863-169A-6 (1-1664)

OY 5 TyrIysThrValIleuLeu 10
Db 1618 TACAAACCGTCTTTTG 1601

RESULT 29
US-08-429-964-6/c
Sequence 6, Application US/08429964
Patent No. 5962243
GENERAL INFORMATION:
APPLICANT: BROWN, MICHAEL S.
APPLICANT: GOLDSTEIN, JOSEPH L.
APPLICANT: REISS, YUVAL
APPLICANT: JAMES, GUY L.
TITLE OF INVENTION: METHODS FOR THE IDENTIFICATION OF FARNESYL
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/429,964
FILING DATE: 27-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/021,625
FILING DATE: 16-FEB-1993
CLASSIFICATION: 435
APPLICATION NUMBER: US 07/822,011
FILING DATE: ABANDONED
CLASSIFICATION: 435
APPLICATION NUMBER: PCT/US/91/02650
FILING DATE: 18-APR-1991
CLASSIFICATION: 435
APPLICATION NUMBER: US 07/615,715
FILING DATE: 20-NOV-1990
CLASSIFICATION: 435
APPLICATION NUMBER: US 07/510,706
FILING DATE: 18-APR-1990 (ABANDONED)

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UTSD:432/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1664 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-429-964-6

Alignment Scores:
Pred. No.: 611
Score: 6.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 10.34%
DB: 2
Matches: 1664
Conservative: 6
Mismatches: 0
Indels: 0
Gaps: 0

US-09-727-892a-99 (1-58) x US-08-429-964-6 (1-1664)

OY 5 TyrlsthrVallLeu10
Db 1618 TACAAACCGTCTTTTG 1601

RESULT 30
US-07-935-087-6/c
Sequence 6, Application US/07935087
Patent No. 6083917
GENERAL INFORMATION:
APPLICANT: BROWN, MICHAEL S.
APPLICANT: GOLDSTEIN, JOSEPH L.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
TITLE OF INVENTION: THE IDENTIFICATION,
TITLE OF INVENTION: CHARACTERIZATION,
TITLE OF INVENTION: AND INHIBITION OF FARNESYL
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1 (converted to ASCII-DOS)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/935,087
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/822,011
FILING DATE: 01/16/92
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UTSD:269/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
TELEX:
INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 1664 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-07-935-087-6

Alignment Scores:
Pred. No.: 611
Score: 6.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 10.34%
DB: 3
Matches: 1664
Conservative: 6
Mismatches: 0
Indels: 0
Gaps: 0

US-09-727-892a-99 (1-58) x US-07-935-087-6 (1-1664)

OY 5 TyrlsthrVallLeu10
Db 1618 TACAAACCGTCTTTTG 1601

RESULT 31
PCT-US93-08062-6/c
Sequence 6, Application PC/TUS9308062
GENERAL INFORMATION:
APPLICANT:
SEQUENCE CHARACTERISTICS: BROWN, MICHAEL S.
SEQUENCE CHARACTERISTICS: GOLDSTEIN, JOSEPH L.
SEQUENCE CHARACTERISTICS: REISS, YUVAL
SEQUENCE CHARACTERISTICS: MARSTERS, JR., JAMES C.
ADDRESSEE: METHODS AND COMPOSITIONS FOR
ADDRESSEE: THE IDENTIFICATION,
ADDRESSEE: CHARACTERIZATION AND
ADDRESSEE: INHIBITION OF
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESSES:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK/ASKII
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08062
FILING DATE: AUGUST 24, 1993
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/935,087
FILING DATE: 24 AUGUST 1992 (24.08.92)
NAME: UNKNOWN
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UTSD377PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
TELEX: NOT APPLICABLE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1664 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US93-08062-6

```
Alignment Scores:
Pred. No.: 611 Length: 1664
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 5 Gaps: 0

US-09-727-892A-99 (1-58) x PCT-US93-08062-6 (1-1664)
QY 5 TyTlYsThVAlLeuLeu 10
Db 1618 TACAAACCGTCTTTTG 1601

RESULT 32
US-09-058-260-13
; Sequence 13, Application US/09058260B
; Patent No. 6218167
; GENERAL INFORMATION:
; APPLICANT: Allen, Larry
; APPLICANT: Aikens, John
; APPLICANT: Fondein, Michael
; APPLICANT: Vonstein, Veronika
; APPLICANT: Demitrijan, David
; APPLICANT: Casadaban, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; FILE REFERENCE: 95-963-H
; CURRENT APPLICATION NUMBER: US/09/058,260B
; EARLIER FILING DATE: 1999-04-10
; EARLIER APPLICATION NUMBER: 60/001,995
; EARLIER FILING DATE: 1996-08-07
; EARLIER APPLICATION NUMBER: 60/009,704
; EARLIER FILING DATE: 1996-01-11
; EARLIER APPLICATION NUMBER: 60/019,580
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: 08/694,078
; EARLIER FILING DATE: 1996-08-08
; EARLIER APPLICATION NUMBER: 08/781,802
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 08/827,810
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1699
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: cloned esterase
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (92)..(1594)
; US-09-058-260-13

Alignment Scores:
Pred. No.: 623 Length: 1699
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 4 Gaps: 0

US-09-727-892A-99 (1-58) x US-09-058-260-13 (1-1699)
QY 40 GluTyAsnLeuPheThr 45
Db 1025 GAGTACAAATTATTAC 1042

RESULT 33
US-09-064-703-4
; Sequence 4, Application US/09064703
; Patent No. 6033894

GENERAL INFORMATION:
; APPLICANT: Craik, Charles S.
; APPLICANT: Unal, Ayce
; APPLICANT: Ganem, Donald E.
; TITLE OF INVENTION: Kaposi's Syndrome Herpesvirus
; TITLE OF INVENTION: Protease and Assembly Protein Compositions and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Denlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/064,703
; FILING DATE: 22-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/044,152
; FILING DATE: 22-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Felthory, Joanne R
; REGISTRATION NUMBER: P42,995
; REFERENCE/DOCKET NUMBER: 2002-0002.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1701 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: KSHV Pr/AP polyprotein including SEQ ID NO:2
; US-09-064-703-4

Alignment Scores:
Pred. No.: 624 Length: 1701
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 3 Gaps: 0

US-09-727-892A-99 (1-58) x US-09-064-703-4 (1-1701)
QY 19 PheProHisGlnIleSer 24
Db 551 TTTCACACCAATTTTCG 568

RESULT 34
US-09-058-260-21
; Sequence 21, Application US/09058260B
; Patent No. 6218167
; GENERAL INFORMATION:
; APPLICANT: Allen, Larry
; APPLICANT: Aikens, John
; APPLICANT: Fondein, Michael
; APPLICANT: Vonstein, Veronika
; APPLICANT: Demitrijan, David
; APPLICANT: Casadaban, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; FILE REFERENCE: 95-963-H
; CURRENT APPLICATION NUMBER: US/09/058,260B
; EARLIER FILING DATE: 1999-04-10
; EARLIER APPLICATION NUMBER: 60/001,995
```

```
; EARLIER FILING DATE: 1996-08-07
; EARLIER APPLICATION NUMBER: 60/009,704
; EARLIER FILING DATE: 1996-01-11
; EARLIER APPLICATION NUMBER: 60/019,580
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: 08/694,078
; EARLIER FILING DATE: 1996-08-08
; EARLIER APPLICATION NUMBER: 08/781,802
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 08/827,810
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentl Ver. 2.0
; SEQ ID NO 21
; LENGTH: 1753
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cloned esterase
; OTHER INFORMATION: gene from bacteria E013
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128)..(1630)
; US-09-058-260-21

Alignment Scores:
Pred. No.: 641 Length: 1753
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
Gaps: 0

US-09-727-892A-99 (1-58) x US-09-058-260-21 (1-1753)
QY 40 GluTyraAsnLeuphethr 45
Db 1061 GAGTACAAATTATTACCA 1078

RESULT 35
; US-09-058-260-31
; Sequence 31, Application US/09058260B
; Patent No. 6218167
; GENERAL INFORMATION:
; APPLICANT: Allen, Larry
; APPLICANT: Aikens, John
; APPLICANT: Foustein, Michael
; APPLICANT: Vonstein, Veronika
; APPLICANT: Demirjian, David
; APPLICANT: Casadaban, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; FILE REFERENCE: 95-963-H
; CURRENT APPLICATION NUMBER: US/09/058,260B
; CURRENT FILING DATE: 1999-04-10
; EARLIER APPLICATION NUMBER: 60/001,995
; EARLIER FILING DATE: 1996-08-07
; EARLIER APPLICATION NUMBER: 60/009,704
; EARLIER FILING DATE: 1996-01-11
; EARLIER APPLICATION NUMBER: 60/019,580
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: 08/694,078
; EARLIER FILING DATE: 1996-08-08
; EARLIER APPLICATION NUMBER: 08/781,802
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 08/827,810
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentl Ver. 2.0
; SEQ ID NO 31
; LENGTH: 1756
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: Description of Artificial Sequence:cloned esterase
; OTHER INFORMATION: gene from bacteria E027
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (130)..(1632)
; US-09-058-260-31

Alignment Scores:
Pred. No.: 642 Length: 1756
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
Gaps: 0

US-09-727-892A-99 (1-58) x US-09-058-260-31 (1-1756)
QY 40 GluTyraAsnLeuphethr 45
Db 1063 GAGTACAAATTATTACCA 1080

RESULT 36
; US-09-058-260-23
; Sequence 23, Application US/09058260B
; Patent No. 6218167
; GENERAL INFORMATION:
; APPLICANT: Allen, Larry
; APPLICANT: Aikens, John
; APPLICANT: Foustein, Michael
; APPLICANT: Vonstein, Veronika
; APPLICANT: Demirjian, David
; APPLICANT: Casadaban, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; FILE REFERENCE: 95-963-H
; CURRENT APPLICATION NUMBER: US/09/058,260B
; CURRENT FILING DATE: 1999-04-10
; EARLIER APPLICATION NUMBER: 60/001,995
; EARLIER FILING DATE: 1996-08-07
; EARLIER APPLICATION NUMBER: 60/009,704
; EARLIER FILING DATE: 1996-01-11
; EARLIER APPLICATION NUMBER: 60/019,580
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: 08/694,078
; EARLIER FILING DATE: 1996-08-08
; EARLIER APPLICATION NUMBER: 08/781,802
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 08/827,810
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentl Ver. 2.0
; SEQ ID NO 23
; LENGTH: 1776
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cloned esterase
; OTHER INFORMATION: gene from bacteria E015
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128)..(1630)
; US-09-058-260-23

Alignment Scores:
Pred. No.: 649 Length: 1776
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
Gaps: 0

US-09-727-892A-99 (1-58) x US-09-058-260-23 (1-1776)
QY 40 GluTyraAsnLeuphethr 45
```


Db 1061 GAGTACAATTATTACA 1078
|||||
RESULT 37
US-08-427-497E-5
; Sequence 5, Application US/08427497E
; Patent No. 5969124
; GENERAL INFORMATION:
; APPLICANT: Lemmon, Vance
; TITLE OF INVENTION: A Method for Characterizing the
; TITLE OF INVENTION: Nucleotide Sequence of L1CAM and
; Patent No. 5969124
; TITLE OF INVENTION: the Nucleotide Sequence
; TITLE OF INVENTION: Characterized Thereby
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fay, Sharpe, Beall, Fagan,
; ADDRESSEE: Minnich & McKee
; STREET: 1100 Superior Avenue
; STREET: Suite 700
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 44114-2518
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; MEDIUM TYPE: storable
; COMPUTER: Compaq Prolinea 5100e
; OPERATING SYSTEM: DOS 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/427,497E
; FILING DATE: April 24, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/904,991
; FILING DATE: June 26, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Minnich, Richard J.
; REGISTRATION NUMBER: 24,175
; REFERENCE/DOCKET NUMBER: CWR 2 149-3-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 861-5582
; TELEFAX: (216) 241-1666
; TELEX: (216) 980162
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1794
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: nucleic acids
; HYPOTHEICAL: irrelevant
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: 17-18 week fetus
; IMMEDIATE SOURCE:
; LIBRARY: StrataGene cDNA Library 936206
; CLONE: 17
; PUBLICATION INFORMATION:
; AUTHORS: Hlavyn, Mary Louise
; AUTHORS: Lemmon, Vance
; TITLE: Molecular structure and functional testing of
; TITLE: human L1CAM: an interspecies comparison.
; JOURNAL: GENOMICS
; VOLUME: 11
; ISSUE:
; PAGES: 416-423
; DATE: 1991
; RELEVANT RESIDUES IN SEQ ID NO: 2731 to 4503
US-08-427-497E-5

Alignment Scores:
Pred. No.: 655 Length: 1794
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 2 Gaps: 0
US-09-727-892a-99 (1-58) x US-08-427-497E-5 (1-1794)
QY 17 GlyHisPheProHisGln 22
|||||
Db 1026 GGCCACTTCCCATCA 1043
RESULT 38
US-09-058-260-3
; Sequence 3, Application US/09058260B
; Patent No. 6218167
; GENERAL INFORMATION:
; APPLICANT: Allen, Larry
; APPLICANT: Aikens, John
; APPLICANT: Fonstein, Michael
; APPLICANT: Vonstein, Veronika
; APPLICANT: Demirjian, David
; APPLICANT: Casadaban, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; FILE REFERENCE: 95-963-H
; CURRENT APPLICATION NUMBER: US/09/058,260B
; CURRENT FILING DATE: 1999-04-10
; EARLIER APPLICATION NUMBER: 60/001,995
; EARLIER FILING DATE: 1996-08-07
; EARLIER APPLICATION NUMBER: 60/009,704
; EARLIER FILING DATE: 1996-01-11
; EARLIER APPLICATION NUMBER: 60/019,580
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: 08/594,078
; EARLIER FILING DATE: 1996-08-08
; EARLIER APPLICATION NUMBER: 08/781,802
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 08/827,810
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 3
; LENGTH: 1896
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: cloned esterase
; OTHER INFORMATION: gene from bacteria E009
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (211)..(1713)
US-09-058-260-3
Alignment Scores:
Pred. No.: 689 Length: 1896
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 4 Gaps: 0
US-09-727-892a-99 (1-58) x US-09-058-260-3 (1-1896)
QY 40 GluTyrAsnLeuPheThr 45
|||||
Db 1144 GAGTACAATTATTACA 1161
RESULT 39
US-09-058-260-17
; Sequence 17, Application US/09058260B
; Patent No. 6218167

```

GENERAL INFORMATION:
APPLICANT: Allen, Larry
APPLICANT: Atkins, John
APPLICANT: Fousteln, Michael
APPLICANT: Vonstein, Veronika
APPLICANT: Demitrijan, David
APPLICANT: Casadaban, Malcolm
TITLE OF INVENTION: Stable Biotocatalysts for Ester Hydrolysis
FILE REFERENCE: 95-963-H
CURRENT APPLICATION NUMBER: US/09/058,260B
CURRENT FILING DATE: 1999-04-10
EARLIER APPLICATION NUMBER: 60/001,995
EARLIER FILING DATE: 1996-08-07
EARLIER APPLICATION NUMBER: 60/009,704
EARLIER FILING DATE: 1996-01-11
EARLIER APPLICATION NUMBER: 60/019,580
EARLIER FILING DATE: 1996-06-12
EARLIER APPLICATION NUMBER: 08/694,078
EARLIER FILING DATE: 1996-08-08
EARLIER APPLICATION NUMBER: 08/781,802
EARLIER FILING DATE: 1997-01-10
EARLIER APPLICATION NUMBER: 08/827,810
EARLIER FILING DATE: 1997-04-11
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 17
LENGTH: 1925
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: cloned esterase
FEATURE:
OTHER INFORMATION: gene from bacteria E008
NAME/KEY: CDS
LOCATION: (127)..(1581)
US-09-058-260-17

Alignment Scores:
Pred. No.: 699 Length: 1925
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
Gaps: 0
DB: 4

US-09-727-892a-99 (1-58) x US-09-058-260-17 (1-1925)

QY 40 GlutyrAsnLeuphethr 45
Db 1012 GAGTACAAATTTATTTACA 1029

RESULT 40
US-08-472-659-30
Sequence 30, Application US/08472659
Patent No. 5831030
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROOKA, No. 5831030uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5831030uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAUCHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
```

```

ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 1950 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
STRAIN: AA31
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1
OTHER INFORMATION: /note="DNA coding for human
OTHER INFORMATION: megakaryocyte differentiation factor."
FEATURE:
NAME/KEY: CDS
LOCATION: 74..1217
US-08-472-659-30

Alignment Scores:
Pred. No.: 707 Length: 1950
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
Gaps: 0
DB: 2

US-09-727-892a-99 (1-58) x US-08-472-659-30 (1-1950)

QY 29 LeutyrAspAlaIysval 34
Db 434 TTATGATGCAAGTG 451

RESULT 41
US-08-474-661-30
Sequence 30, Application US/08474661
Patent No. 5874253
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROOKA, No. 5874253uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5874253uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAUCHI, Kozo
```

APPLICANT: YAMAGUCHI, No. 5874253coml
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,661
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANER
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 1950 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
STRAIN: A431
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1
OTHER INFORMATION: /note="DNA coding for human
OTHER INFORMATION: megakaryocyte differentiation factor."
FEATURE:
NAME/KEY: CDS
LOCATION: 74..1217
US-08-474-661-30
Alignment Scores:
Pred. No.: 707 Length: 1950
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 2 Gaps: 0
US-09-727-892a-99 (1-58) x US-08-474-661-30 (1-1950)
QY 29 LeuTyrAspAlaIysVal 34
Db 434 TTATACGATGCAAGTG 451
RESULT 42
US-08-611-977-30
Sequence 30, Application US/08611977
Patent No. 5972886
GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURODOKA, No. 5972886uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, No. 5972886uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, No. 5972886coml
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,977
FILING DATE: 06-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 1950 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
STRAIN: A431
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1
OTHER INFORMATION: /note="DNA coding for human
OTHER INFORMATION: megakaryocyte differentiation factor."
FEATURE:
NAME/KEY: CDS
LOCATION: 74..1217
US-08-611-977-30
Alignment Scores:
Pred. No.: 707 Length: 1950
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 2 Gaps: 0
US-09-727-892a-99 (1-58) x US-08-611-977-30 (1-1950)

QY 29 LeuTYTAspAlaLysVal 34
|||||
Db 434 TTATACGATCCTCAAGTG 451

RESULT 43
US-09-058-260-5
; Sequence 5, Application US/09058260B
; Patent No. 6218167
; GENERAL INFORMATION:
; APPLICANT: Allen, Larry
; APPLICANT: Aikens, John
; APPLICANT: Feinstein, Michael
; APPLICANT: Vonstein, Veronika
; APPLICANT: Demitrijan, David
; APPLICANT: Casadaban, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; FILE REFERENCE: 95-963-H
; CURRENT APPLICATION NUMBER: US/09/058,260B
; EARLIER FILING DATE: 1999-04-10
; EARLIER APPLICATION NUMBER: 60/001,995
; EARLIER FILING DATE: 1996-08-07
; EARLIER APPLICATION NUMBER: 60/009,704
; EARLIER FILING DATE: 1996-01-11
; EARLIER APPLICATION NUMBER: 60/019,580
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: 08/694,078
; EARLIER FILING DATE: 1996-08-08
; EARLIER APPLICATION NUMBER: 08/781,802
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 08/827,810
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1952
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cloned esterase
; FEATURE:
; OTHER INFORMATION: gene from bacteria E011
; NAME/KEY: CDS
; LOCATION: (197)..(1699)
US-09-058-260-5

Alignment Scores:
Pred. No.: 708 Length: 1952
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
Gaps: 0

US-09-727-892A-99 (1-58) x US-09-058-260-5 (1-1952)

QY 40 GluTYrAsnLeuPheThr 45
|||||
Db 1130 GAGTACAATTATTATTACA 1147

RESULT 44
US-09-058-260-19
; Sequence 19, Application US/09058260B
; Patent No. 6218167
; GENERAL INFORMATION:
; APPLICANT: Allen, Larry
; APPLICANT: Aikens, John
; APPLICANT: Feinstein, Michael
; APPLICANT: Vonstein, Veronika
; APPLICANT: Demitrijan, David
; APPLICANT: Casadaban, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; FILE REFERENCE: 95-963-H
; CURRENT APPLICATION NUMBER: US/09/058,260B

CURRENT FILING DATE: 1999-04-10
EARLIER APPLICATION NUMBER: 60/001,995
EARLIER FILING DATE: 1996-08-07
EARLIER APPLICATION NUMBER: 60/009,704
EARLIER FILING DATE: 1996-01-11
EARLIER APPLICATION NUMBER: 60/019,580
EARLIER FILING DATE: 1996-06-12
EARLIER APPLICATION NUMBER: 08/694,078
EARLIER FILING DATE: 1996-08-08
EARLIER APPLICATION NUMBER: 08/781,802
EARLIER FILING DATE: 1997-01-10
EARLIER APPLICATION NUMBER: 08/827,810
EARLIER FILING DATE: 1997-04-11
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 19
LENGTH: 1957
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:cloned esterase
OTHER INFORMATION: gene from bacteria E010
FEATURE:
NAME/KEY: CDS
LOCATION: (88)..(1590)
US-09-058-260-19

Alignment Scores:
Pred. No.: 709 Length: 1957
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
Gaps: 0

US-09-727-892A-99 (1-58) x US-09-058-260-19 (1-1957)

QY 40 GluTYrAsnLeuPheThr 45
|||||
Db 1021 GAGTACAATTATTATTACA 1038

RESULT 45
US-08-946-026-15
; Sequence 15, Application US/08946026
; Patent No. 6034218
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Dillon, David G.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Mitcham, Jennifer L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; TITLE OF INVENTION: AND IMMUNODIAGNOSIS OF PROSTATE CANCER
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/946,026
; FILING DATE: 07-OCT-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.424C1

TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1996 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-946-026-15

Alignment Scores:
Pred. No.: 722 Length: 1996
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.34% Indels: 0
DB: 3 Gaps: 0

US-09-727-892A-99 (1-58) x US-08-946-026-15 (1-1996)

OY 7 ThrValLeuLeuTYRCys 12
|||||
Db 1301 ACAGTTTGCTACTACTGT 1318

Search completed: November 5, 2002, 04:24:00
Job time : 96 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 04:56:07 : Search time 26 Seconds
(without alignments)
86.374 Million cell updates/sec

Title: US-09-727-892a-99

Perfect score: 58
Sequence: 1 MERKKTIVLYCDEIKGHP.....YENLFTKRYAIIIEIKEL 58

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt.40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	7	12.1	348	1	GPDA_RABIT
2	6	10.3	186	1	NADM_PYRAB
3	6	10.3	186	1	NADM_PYRHO
4	6	10.3	211	1	DER1_YEAST
5	6	10.3	216	1	YVU4_YEAST
6	6	10.3	232	1	6FGL_CAOGR
7	6	10.3	237	1	KRYA_BACSU
8	6	10.3	273	1	ROCI_NICSY
9	6	10.3	279	1	ROCI_NICPL
10	6	10.3	291	1	ROCI_NICSY
11	6	10.3	292	1	ROCI_NICPL
12	6	10.3	320	1	HEW2_SALTY
13	6	10.3	364	1	PARB_HUMAN
14	6	10.3	372	1	PARA_HUMAN
15	6	10.3	372	1	PARA_MOUSE
16	6	10.3	372	1	PARA_MOUSE
17	6	10.3	378	1	PARA_RAT
18	6	10.3	380	1	SPB7_HUMAN
19	6	10.3	388	1	SPB7_HUMAN
20	6	10.3	421	1	YB1U_ECOLI
21	6	10.3	461	1	V218_FOWPV
22	6	10.3	503	1	SECD_HELPY
23	6	10.3	509	1	SECD_HELPY
24	6	10.3	513	1	YB85_METUA
25	6	10.3	518	1	YB95_METUA
26	6	10.3	526	1	SECD_HELPY
27	6	10.3	545	1	HGT1_CANAL
28	6	10.3	550	1	SVR_CORGL
29	6	10.3	649	1	VAY3_SCHPO
30	6	10.3	682	1	ALT_BPT4
31	6	10.3	698	1	ALT_BPT2
32	6	10.3	698	1	ALT_BPT6
33	6	10.3	719	1	TE80_TETTH

34	6	10.3	789	1	CAD6_RAT	P55280 rattus norv
35	6	10.3	789	1	CAD9_HUMAN	O9ulb4 homo sapien
36	6	10.3	790	1	CAD6_CHICK	O90762 gallus gall
37	6	10.3	790	1	CAD6_HUMAN	P55285 homo sapien
38	6	10.3	790	1	CAD6_MOUSE	P97326 homo sapien
39	6	10.3	841	1	CHS1_PHYBL	P87073 phycomyces
40	6	10.3	850	1	PRTP_HCMVA	P16724 human cytom
41	6	10.3	858	1	CHS1_RH1OL	P30594 rhizopus ol
42	6	10.3	919	1	SVL_THEMEA	P46213 thermotoga
43	6	10.3	945	1	AMPE_MOUSE	P16406 mus musculu
44	6	10.3	957	1	AMPE_HUMAN	O07075 homo sapien
45	6	10.3	1015	1	PDNG_ECOLI	P24183 escherichia

ALIGNMENTS

RESULT 1				
ID	GPDA_RABIT	STANDARD:	PRT:	348 AA.
AC	P08507:			
DT	01-AUG-1988 (Rel. 08, Created)			
DT	01-AUG-1988 (Rel. 08, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Glycerol-3-phosphate dehydrogenase [NAD+], cytoplasmic (EC 1.1.1.8)			
DE	(GPD-C) (GPDH-C).			
GN	GPD.			
OS	Oryctolagus cuniculus (Rabbit)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
OX	NCBI_TaxID=9986;			
RN	[1]			
RP	SECONDARY STRUCTURE PREDICTION.			
RX	MEDLINE=81003924; PubMed=6773774;			
RA	Otto J., Argos P., Rossmann M.G.;			
RT	"Prediction of secondary structural elements in glycerol-3-phosphate dehydrogenase by comparison with other dehydrogenases.";			
RL	Eur. J. Biochem. 109:325-330(1980).			
CC	-1- CATALYTIC ACTIVITY: Sn-glycerol 3-phosphate + NAD(+) = glycerone phosphate + NADH.			
CC	-1- SUBUNIT: HOMODIMER.			
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic.			
CC	-1- SIMILARITY: BELONGS TO THE NAD-DEPENDENT GLYCEROL-3-PHOSPHATE DEHYDROGENASE FAMILY.			
DR	PIR: A32512; A32512.			
DR	InterPro: IPR001652; NAD_Gly3P_dh.			
DR	PIfam: PF01210; NAD_Gly3P_dh: 1.			
DR	PRINTS: PR00077; GPDHGRNASE.			
DR	ProDom: PD001649; NAD_Gly3P_dh: 1.			
DR	PROSITE: PS00957; NAD_G3PDH: 1.			
KW	Oxidoreductase; NAD.			
FT	INIT_MET 0			
SO	SEQUENCE 348 AA; 37478 MW; 74386ED5E2C60E45 CRC64;			
Query Match 12.1%; Score 7; DB 1; Length 348;				
Best Local Similarity 100.0%; Pred. No. 2.2;				
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	12 CDEIKGH 18			
DB	101 CDEIKGH 107			
RESULT 2				
ID	NADM_PYRAB	STANDARD:	PRT:	186 AA.
AC	O9UYD4;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Nicotinamide-nucleotide adenyllyltransferase (EC 2.7.7.1) (NAD(+))			
DE	pyrophosphorylase) (NAD(+))			
GN	PAB1318.			

```

OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORSAY;
RA Helling R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
  structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + nicotinamide ribonucleotide =
  diphosphate + NAD(+).
CC -1- PATHWAY: NAD BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ARCHAEL NMN ADENYLYLTRANSFERASE
  FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ248288; CAB50478.1; -.
DR HSSP; 026253; IEJ2.
DR InterPro: IPR001994; Cytidylyltransf.
DR Pfam; PF01467; Cytidylyltransf. 1.
DR Transferrase; Nucleotidyltransferase; NAD; Complete proteome.
KW SEQUENCE 186 AA; 21417 MW; E8230B688481386E CRC64;
SQ
Query Match 10.3%; Score 6; DB 1; Length 186;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 53 EYKEI 58
Db 152 EYKEI 157

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP000002; BAA29550.1; -.
DR HSSP; 026253; IEJ2.
DR InterPro: IPR001994; Cytidylyltransf.
DR Pfam; PF01467; Cytidylyltransf. 1.
DR Transferrase; Nucleotidyltransferase; NAD; Complete proteome.
KW SEQUENCE 186 AA; 21392 MW; ICE5A40C84ADB34 CRC64;
SQ
Query Match 10.3%; Score 6; DB 1; Length 186;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 53 EYKEI 58
Db 152 EYKEI 157

```

```

RESULT 3
ID NADM_PYRHO STANDARD; PRT; 186 AA.
AC 058211;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nicotinamide-nucleotide adenylyltransferase (EC 2.7.7.1) (NAD(+)-
  pyrophosphorylase) (NAD(+)-diphosphorylase) (NMN adenylyltransferase).
GN P00464.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OTS3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,
  Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
  Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohuku Y.,
  Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
  Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
  Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
  thermophilic archaeobacterium, Pyrococcus horikoshii OTS3.";
RL DNA Res. 5:55-76(1998).
CC -1- CATALYTIC ACTIVITY: ATP + nicotinamide ribonucleotide =
  diphosphate + NAD(+).
CC -1- PATHWAY: NAD BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ARCHAEL NMN ADENYLYLTRANSFERASE
  FAMILY.

```

```

RESULT 4
ID DERL_YEAST STANDARD; PRT; 211 AA.
AC P38307;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DERL protein.
GN DERL OR YBR201W OR YBR1413.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
  Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=9506384; PubMed=7975899;
RA Mallet L., Bussereau F., Jaquet M.;
RT "Nucleotide sequence analysis of an 11.7 kb fragment of yeast
  chromosome II including BEM1, a new gene of the WD-40 repeat family
  and a new member of the KRE2/MNT1 family.";
RL Yeast 10:819-831(1994).
RN [2]
RP REVISIONS, AND CHARACTERIZATION.
RX MEDLINE=96181354; PubMed=8631297;
RA Knop M., Finger A., Braun T., Hellmuth K., Wolf D.H.;
RT "Derl, a novel protein specifically required for endoplasmic
  reticulum degradation in yeast.";
RL EMO J. 15:753-763(1996).
CC -1- FUNCTION: SPECIFICALLY REQUIRED FOR THE DEGRADATION PROCESS OF
  MISFOLEDDED ENDOPLASMIC RETICULUM LUMINAL PROTEINS. COULD ACT AS A
  COMPONENT OF THE SUBSTRATE-RECOGNIZING SYSTEM, BUT IT COULD ALSO
  ACT IN MECHANISMS THAT ARE INVOLVED IN CORRECT LOCALIZATION OR
  FUNCTION OF THE PROTEASE(S).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
  FRAMESHIFTS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z21487; CAA79688.1; ALT_FRAME.
DR EMBL; Z36070; CAA85165.1; -.

```


DR EMBL: Z36069; CAA85164.1; -
DR EMBL: X92435; CAA63165.1; -
DR PIR: S45450; S45450.
DR PIR: S34026; S34026.
DR SGD: S0000405; DER1.
KW Transmembrane; Endoplasmic reticulum.
FT TRANSMEM 5 25 POTENTIAL.
FT TRANSMEM 56 76 POTENTIAL.
FT TRANSMEM 94 114 POTENTIAL.
FT TRANSMEM 143 163 POTENTIAL.
FT CONFLICT 145 145 V -> A (IN REF. 1).
SQ SEQUENCE 211 AA; 24419 MW; 41511696F93360D CRC64;

Query Match 10.3%; Score 6; DB 1; Length 211;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 KVVYSY 38
| | | | |
DB 39 KVVYSY 44

RESULT 5
ID Y104_YEAST STANDARD; PRT; 216 AA.
Y104_YEAST
AC P40576;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Hypothetical 24.6 kDa protein in DAL81-YVH1 intergenic region.
GN Y1024C.
OS Saccharomyces cerevisiae (Baker's yeast).
CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SZ88C / AB972;
RA Barrall B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagsels K., Jones M.,
RA Louis E., Lye G., Moulé S., Moulé T., Odell C., Pearson D.,
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z38061; CAA86184.1; -
DR PIR: S48486; S48486.
DR SGD: S0001463; GIP1.
KW Hypothetical protein.
SQ SEQUENCE 216 AA; 24597 MW; E63342B8D5DD079 CRC64;

Query Match 10.3%; Score 6; DB 1; Length 216;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MERRYK 6
| | | | |
DB 71 MERRYK 76

RESULT 6
ID 6PGL_CAUCR STANDARD; PRT; 232 AA.
AC 09A6N1;
RT

DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 6-phosphogluconolactonase (EC 3.1.1.31) (6PGL).
GN PGL OR CC2056.
OS Caulobacter crescentus.
CC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter
OX NCBI_TaxID=69394;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nielsen W.C., Feldblyum T.V., Iaub M.T., Paulsen I.T., Nelson K.E.,
RA Eelsen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT *Complete genome sequence of Caulobacter crescentus.*;
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -1- FUNCTION: HYDROLYSIS OF 6-PHOSPHOGLUCONOLACTONE TO 6-
CC PHOSPHOGLUCONATE.
CC -1- CATALYTIC ACTIVITY: 6-PHOSPHO-D-GLUCONO-1,5-LACTONE + H(2)O = 6-
CC PHOSPHO-D-GLUCONATE.
CC -1- PATHWAY: SECOND STEP IN PENTOSE PHOSPHATE PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE GLUCOSAMINE/GALACTOSAMINE-6-PHOSPHATE
CC ISOMERASE FAMILY. 6-PHOSPHOGLUCONOLACTONASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE005879; AM24029.1; -
DR TIGR: CC2056; -
DR InterPro: IPR000457; Glucosamine_Iso.
DR Pfam: PF01182; Glucosamine_Iso; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 232 AA; 24233 MW; 8FBEA52F720F87B6 CRC64;

Query Match 10.3%; Score 6; DB 1; Length 232;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 EDLYDA 32
| | | | |
DB 15 EDLYDA 20

RESULT 7
ID KYKA_BACSU STANDARD; PRT; 237 AA.
AC P21884;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 27.3 kDa protein in PDHA 5' region (ORF5).
GN KYKA.
OS Bacillus subtilis.
CC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=90368558; PubMed=1697575;
RA Hemila H., Palva A., Paulin L., Arvidson S., Palva I.;
RT "Secretory S complex of Bacillus subtilis: sequence analysis and

```

RT Identity to pyruvate dehydrogenase.";
RL J. Bacteriol. 172:5052-5063(1990).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: M57435; AAA62680.1; ALT_INIT.
DR Subtilisin: Bg10206; YKXA.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 237 AA; 27322 MW; D6924F6CB2ED724B CRC64;

Query Match
Best Local Similarity 10.3%; Score 6; DB 1; Length 237;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 27 EDLYDA 32
Db 224 EDLYDA 229

RESULT 8
ROCL_NICSY STANDARD; PRT; 273 AA.
ID ROCL_NICSY
AC 008935;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 29 kDa ribonucleoprotein A, chloroplast precursor (CP29A).
OS Nicotiana sylvestris (Wood tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
ON NCBI_TaxID=4096;
RX MEDLINE=92093607; PubMed=1721701;
RA Ye L., Li Y., Fukami-Kobayashi F., Go M., Konishi T., Matanbe A.,
RA Suglura M.;
RT "Diversity of a ribonucleoprotein family in tobacco chloroplasts: two
RT new chloroplast ribonucleoproteins and a phylogenetic tree of ten
RT chloroplast RNA-binding domains."
RL Nucleic Acids Res. 19:6485-6490(1991).
CC -1- FUNCTION: COULD BE INVOLVED IN SPLICING AND/OR PROCESSING OF
CC CHLOROPLAST RNA'S.
CC -1- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X61113; CAA43427.1; -.
DR HSSP: P19339; ISXL.
DR Mendel: 15244; NicSY:2406;15244.
DR InterPro: IPR000504; RRM.
DR Pfam: PF00076; rrm; 2.
DR SMART: SM00360; RRM; 2.
DR PROSITE: PS00102; RRM; 2.
DR PROSITE: PS00030; RRM_RNP_1; 2.
DR RNA-binding; Ribonucleoprotein; Repeat; mRNA processing; Chloroplast;
KW Transit peptide.
FT CHAIN 1 58 CHLOROPLAST (POTENTIAL).
FT TRANSIT 1 58 CHLOROPLAST (POTENTIAL).
FT CHAIN 59 273 29 KDA RIBONUCLEOPROTEIN A.

```

```

FT DOMAIN 43 49 POLY-SER.
FT DOMAIN 59 82 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 87 165 RNA-BINDING (RRM) 1.
FT DOMAIN 166 187 LINKER (GLY-RICH).
FT DOMAIN 188 266 RNA-BINDING (RRM) 2.
SQ SEQUENCE 273 AA; 29674 MW; F45ADF81B3BF133 CRC64;

Query Match
Best Local Similarity 10.3%; Score 6; DB 1; Length 273;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 DAKVYV 36
Db 216 DAKVYV 221

RESULT 9
ROCL_NICPL STANDARD; PRT; 279 AA.
ID ROCL_NICPL
AC P49313;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 30 kDa ribonucleoprotein, chloroplast precursor (CP-RBP30).
OS Nicotiana plumbaginifolia (Leadwort-leaved tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
ON NCBI_TaxID=4092;
RX MEDLINE=93024313; PubMed=1406585;
RA Mieszkaz M., Klahre U., Levy J.H., Goodall G.J., Filipowicz W.;
RT "Multiple plant RNA binding proteins identified by PCR: expression of
RT cDNAs encoding RNA binding proteins targeted to chloroplasts in
RT Nicotiana plumbaginifolia."
RL Mol. Gen. Genet. 234:390-400(1992).
CC -1- FUNCTION: COULD BE INVOLVED IN SPLICING AND/OR PROCESSING OF
CC CHLOROPLAST RNA'S.
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- TISSUE SPECIFICITY: IT IS EXPRESSED AT HIGH LEVELS IN THE LEAVES
CC AND SEEDLINGS, AND LOWER LEVELS ARE SEEN IN THE STEMS AND ROOTS.
CC -1- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X65118; CAA46234.1; -.
DR HSSP: P19339; ISXL.
DR InterPro: IPR000504; RRM.
DR Pfam: PF00076; rrm; 2.
DR SMART: SM00360; RRM; 2.
DR PROSITE: PS00102; RRM; 2.
DR PROSITE: PS00030; RRM_RNP_1; 2.
DR RNA-binding; Ribonucleoprotein; Repeat; mRNA processing; Chloroplast;
KW Transit peptide.
FT CHAIN 1 279 CHLOROPLAST (POTENTIAL).
FT TRANSIT 1 279 30 KDA RIBONUCLEOPROTEIN.
FT DOMAIN 87 165 RNA-BINDING (RRM) 1.
FT DOMAIN 166 193 LINKER (GLY-RICH).
FT DOMAIN 194 272 RNA-BINDING (RRM) 2.
SQ SEQUENCE 279 AA; 30461 MW; 827FDAB5B3FDEFD32 CRC64;

Query Match
Best Local Similarity 10.3%; Score 6; DB 1; Length 279;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 31 DAKVY 36
 DB 222 DAKVY 227

RESULT 10 ROC2_NICSY

STANDARD: PRT: 291 AA.
 ID ROC2_NICSY 008937;
 AC 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 OS 29 kDa ribonucleoprotein B, chloroplast precursor (CP29B).
 OC Nicotiana sylvestris (Wood tobacco).
 CC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 CC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots;
 CC Asteridae: euasterids I: Solanales: Solanaceae; Nicotiana.
 OX NCBI_TaxID=4096;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RX MEDLINE=92093607; PubMed=1721701;
 RA Ye L., Li Y., Fukami-Kodayashi F., Go M., Konishi T., Watanabe A.,
 RA Sugitara M.;
 RT "Diversity of a ribonucleoprotein family in tobacco chloroplasts: two
 RT new chloroplast RNA-binding domains";
 RT chloroplast RNA-binding domains";
 RL Nucleic Acids Res. 19:6485-6490(1991).
 CC -1- FUNCTION: COULD BE INVOLVED IN SPLICING AND/OR PROCESSING OF
 CC CHLOROPLAST RNA S.
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: X61114; CAA43428.1; -;
 CC DR HSSP: P19339; ISXL.
 CC DR Mendel: 15245; Nicsy:2406;15245.
 CC DR InterPro: IPR000504; RRM.
 CC DR Pfam: PF00076; rrm; 2.
 CC DR SMART: SM00360; RRM; 2.
 CC DR PROSITE: PS50102; RRM; 2.
 CC DR PROSITE: PS00030; RRM_NRP_1; 2.
 CC KW RNA-binding; Ribonucleoprotein; Repeat; mRNA processing; Chloroplast;
 CC KW Transist peptide.
 CC FT TRANSIT 1 ?
 CC FT CHAIN ? 291 CHLOROPLAST (POTENTIAL).
 CC FT DOMAIN 87 165 29 KDA RIBONUCLEOPROTEIN B.
 CC FT DOMAIN 166 206 LINKER (GLY-RICH) 1.
 CC FT DOMAIN 207 285 RNA-BINDING (RRM) 2.
 CC SQ SEQUENCE 291 AA; 31114 MW; 7DF5F86D17E151F4 CAC64;
 CC
 CC Query Match 10.3%; Score 6; DB 1; Length 291;
 CC Best Local Similarity 100.0%; Pred. No. 24;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC OY 31 DAKVY 36
 CC DB 235 DAKVY 240

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 31 kDa ribonucleoprotein, chloroplast precursor (CP-RBP31).
 OS Nicotiana plumbaginifolia (Leadwort-leaved tobacco).
 CC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
 CC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots;
 CC Asteridae: euasterids I: Solanales: Solanaceae; Nicotiana.
 OX NCBI_TaxID=4092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RX MEDLINE=93024312; PubMed=1406585;
 RA Mieszcak M., Klahre U., Levy J.H., Goodall G.J., Filipowicz W.;
 RT "Multiple plant RNA binding proteins identified by PCR: expression of
 RT cDNAs encoding RNA binding proteins targeted to chloroplasts in
 RT Nicotiana plumbaginifolia";
 RL Mol. Gen. Genet. 234:390-400(1992).
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- TISSUE SPECIFICITY: IT IS EXPRESSED AT HIGH LEVELS IN THE LEAVES
 CC AND SEEDLINGS, AND LOWER LEVELS ARE SEEN IN THE STEMS AND ROOTS.
 CC -1- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: X65117; CAA46233.1; -;
 CC DR HSSP: P19339; ISXL.
 CC DR InterPro: IPR000504; RRM.
 CC DR Pfam: PF00076; rrm; 2.
 CC DR SMART: SM00360; RRM; 2.
 CC DR PROSITE: PS50102; RRM; 2.
 CC DR PROSITE: PS00030; RRM_NRP_1; 2.
 CC KW RNA-binding; Ribonucleoprotein; Repeat; mRNA processing; Chloroplast;
 CC KW Transist peptide.
 CC FT TRANSIT 1 ?
 CC FT CHAIN ? 292 CHLOROPLAST (POTENTIAL).
 CC FT DOMAIN 88 166 31 KDA RIBONUCLEOPROTEIN.
 CC FT DOMAIN 167 207 LINKER (GLY-RICH) 1.
 CC FT DOMAIN 208 286 RNA-BINDING (RRM) 2.
 CC SQ SEQUENCE 292 AA; 31219 MW; 30A202C55E3E791 CAC64;
 CC
 CC Query Match 10.3%; Score 6; DB 1; Length 292;
 CC Best Local Similarity 100.0%; Pred. No. 24;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC OY 31 DAKVY 36
 CC DB 236 DAKVY 241

RESULT 12
 HEMZ_SALTY
 ID HEMZ_SALTY STANDARD: PRT: 320 AA.
 AC P37408;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Ferriohelataase (EC 4.99.1.1) (Prothomene ferro-lyase) (Heme
 DE synthetase).
 GN HEMH OR VISA OR STM0489.
 OS Salmonella typhimurium.
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC OC Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA Mclelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,

RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Milvaney E.,
 RA Rran E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 RT LT2.",
 RL Nature 413:852-856(2001).
 RN
 RP SEQUENCE OF 1-168 FROM N.A.
 RC STRAIN-LT2;
 RX MEDLINE=95113777; PubMed=7814329;
 RA Guiterrez J.A., Gsconka L.N.;
 RT "Isolation and characterization of adenylate kinase (ack) mutations
 RT in *Salmonella typhimurium* which block the ability of glycine betaine
 RT to function as an osmoprotectant.";
 RL Bacteriol. 177:390-400(1995).
 CC -1- FUNCTION: CATALYZES THE FERROUS INSERTION INTO PROTOPORPHYRIN IX.
 CC -1- CATALYTIC ACTIVITY: Protoporphyrin + Fe(2+) = protoheme + 2 H(+).
 CC -1- PATHWAY: LAST STEP IN PROTOHEME BIOSYNTHESIS.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE FERROCHELATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE008718; AL19443.1; -
 DR EMBL; L26246; AAA65970.1; -
 DR StGene; SG10151; hemH.
 DR Interpro: IPR001015; Ferriochelataase.
 DR Pfam: PF00762; Ferriochelataase; 1.
 DR PROSITE; PS00534; FERROCHELATASE; 1.
 DR Porphyrin biosynthesis; Heme biosynthesis; Lyase; Iron;
 KW Complete proteome.
 FT METAL 194 194 IRON (BY SIMILARITY).
 FT METAL 275 275 IRON (BY SIMILARITY).
 FT CONFLICT 41 41 P -> S (IN REF. 2).
 FT CONFLICT 82 82 E -> Q (IN REF. 2).
 FT CONFLICT 104 104 S -> C (IN REF. 2).
 FT CONFLICT 149 149 R -> P (IN REF. 2).
 FT CONFLICT 159 159 I -> M (IN REF. 2).
 SQ SEQUENCE 320 AA; 35914 MM; F391D1376239817 CRC64;
 Query Match 10.3%; Score 6; DB 1; Length 320;
 Best Local Similarity 100.0%; Pred. NO. 26;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 46 KKYAYI 51
 Db 294 KKYAYI 299
 RESULT 13
 PARB HUMAN STANDARD; PRU: 364 AA.
 AC Q9HBT1; Q9NSP7; Q9UGT3; Q9Y3L6; Q9Y3L7; Q9Y368;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Beta-parvin (Affixin) (CGI-56).
 GN PARVB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21102127; PubMed=11171322;

RA Olaski T.M., Noegel A.A., Korenbaum E.;
 RT "Parvin, a 42 kDa focal adhesion protein, related to the alpha-actinin
 RT superfamily".
 RL J. Cell Sci. 114:525-538(2001).
 RN
 RP SEQUENCE FROM N.A.
 RA Yamaji S., Suzuki A., Sugiyama Y., Koide Y., Yoshida M., Kanamori H.,
 RA Mouri H., Ohno S., Ishigatsubo Y.;
 RT "A novel ILK binding protein, affixin, is involved in the early stage
 RT of cell-substrate interaction.";
 RL J. Cell Biol. 0:0-0(2001).
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,
 RA Clamp M., Smink L.J., Altschough R., Almeida J.P., Babage A.,
 RA Baguley C., Bailey J., Barlow K., Bates K.N., Beasley O., Bird C.P.,
 RA Blakey S., Bridgeman A.M., Buck D., Burgess J., Burrill W.D.,
 RA Burton J., Carder C., Carter N.P., Chen Y., Clark G., Clegg S.M.,
 RA Cobley V., Cole C.G., Collier R.E., Connor R.E., Conroy D., Corby N.,
 RA Cowley G.J., Cox A.V., Davis J., Dawson E., Dhani P.D., Dockree C.,
 RA Dodsworth S.J., Dublin R.M., Ellington A., Evans K.L., Fey J.M.,
 RA Fleming K., French L., Garner A.A., Gilbert J.G.R., Goward M.E.,
 RA Griffiths M.N., Hall C., Hall R., Hall T., Hall T., Kershaw J.,
 RA Heathcote R.W., Ho S., Holmes S., Hunt S.E., Jones M.C., Kershaw J.,
 RA Kimberley A., King A., Laird G.K., Langford C.F., Leverish M.A.,
 RA Lloyd C., Lloyd D.M., Martyn I.D., Mashreghi-Mohammadi M.,
 RA Matthews L., Mccann O.T., Mcclay J., McLaren S., Mcmurray A.A.,
 RA Milne S.A., Mortimore B.J., Odell C.N., Pavitt R., Pearce A.V.,
 RA Pearson D., Phillimore B.J., Phillips S.H., Plumb R.W., Ramsey H.,
 RA Ramsey Y., Rogers L., Ross M.T., Scott C.E., Senra H.K., Skuce C.D.,
 RA Smalley S., Smith M.L., Soderlund C., Spragon L., Steward C.A.,
 RA Sulston J.E., Swann R.M., Vaudin M., Wall M., Wallis J.M.,
 RA Whiteley M.N., Whiteley D., Williams L., Williams S., Williamson H.,
 RA Wilmer T.E., Wilming L., Wright C.L., Hubbard T., Bentley D.R.,
 RA Beck S., Rogers J., Shimizu N., Minoshima S., Kawasaki K., Sasaki T.,
 RA Asakawa S., Kudoh J., Shintani A., Shibuya K., Yoshizaki Y., Aoki N.,
 RA Mitsuyama S., Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S.,
 RA Do A., Do T., Dorman A., Fang F., Fu Y., Hua A., Kenton S.,
 RA Lai H., Lao H.I., Lewis S., Lin S.-P., Loh P., Malai E.,
 RA Nguyen T., Pan H., Phan S., Qi S., Qian Y., Ray L., Ren O., Shaull S.,
 RA Sloan D., Song L., Wang Q., Wang Y., Wang Z., White J., Williamson D.,
 RA Wu H., Yao Z., Zhan M., Zhang G., Chisoe S., Murray J., Miller N.,
 RA Mink P., Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H.,
 RA Bourne S., Cordes M., Du Z., Fulton L., Goela D., Graves T.,
 RA Hawkins J., Hinds K., Kemp K., Latreille P., Layman D., Ozerky P.,
 RA Rohlfing T., Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K.,
 RA Nelson J., Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R.,
 RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saita S.,
 RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
 RA Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
 RA Peyrard M., Kedra D., Seroussi E., Eransson I., Tapia I., Bruder C.E.,
 RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
 RA Khan A.S., Lane L., Tiliann Y., Wright H.;
 RT "The DNA sequence of human chromosome 22.";
 RL Nature 402:489-495(1999).
 RN
 RP SEQUENCE OF 15-364 FROM N.A.
 RA Goward M.E., Huckle E.J.;
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Probably plays a role in the regulation of cell adhesion
 CC and cytoskeleton organization.
 CC -1- SUBUNIT: Interacts with integrin-linked protein kinase and actin.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; localized to focal adhesions.
 CC -1- SIMILARITY: BELONGS TO THE PARVIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.

CC -1- CAUTION: Ref.3 sequence differs from that shown due to a
CC number of frameshifts.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL; AF237769; BAB62077.1; -
CC EMBL; AB048276; BAB62077.1; -
CC EMBL; AF151814; AAD34051.1; ALT_FRAME.
CC EMBL; AL031595; CAB63068.1; -
CC EMBL; AL033543; CAB42846.1; ALT_SEQ.
CC EMBL; AL033543; CAB42845.1; ALT_SEQ.
CC EMBL; AL159142; CAB76900.1; -
CC InterPro: IPR001715; Calponin_hom.
CC SMART: SM00033; CH; 2.
CC PROSITE: PS50021; CH; 2.
CC Cell adhesion; Cytoskeleton; Actin-binding; Repeat.
CC DOMAIN 87 194 CH 1.
CC FT DOMAIN 254 361 CH 2.
CC FT CONFLICT 349 349 T->P (IN REF. 3).
CC SEQUENCE 364 AA; 41714 MW; 4BA4B50C83083DC7 CRC64;
OY 41 YNLFTK 46 10.3%; Score 6; DB 1; Length 364;
Db 354 YNLFTK 359 Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 14
PARA_HUMAN STANDARD: PRT; 372 AA.
AC Q9NVD7; Q96C85;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Alpha-parvin (Calponin-like integrin-linked kinase binding protein)
DE (CH-ILKBP).
GN PARVA.
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21102127; PubMed=11171322;
RA Olaki T.M., Noegel A.A., Korenbaum E.;
RT "Parvin, a 42 kDa focal adhesion protein, related to the alpha-actinin
RT superfamily";
RL J. Cell Sci. 114:525-538(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21229705; PubMed=11331308;
RA Tu Y., Huang Y., Zhang Y., Hua Y., Wu C.;
RT "A new focal adhesion protein that interacts with integrin-linked
RT kinase and regulates cell adhesion and spreading";
RL J. Cell Biol. 153:585-598(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX Isegal T., Oca T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahara K., Masuno Y., Sasaki N.;
RT "NDO human cDNA sequencing project";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [4]

RP SEQUENCE FROM N.A.
RC Tissue-Brain;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Probably plays a role in the regulation of cell adhesion
CC and cytoskeleton organization.
CC -1- SUBUNIT: Interacts with integrin-linked protein kinase, actin and
CC paxillin LD1 and LD4 motifs.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic; localized to focal adhesions.
CC -1- SIMILARITY: BELONGS TO THE PARVIN FAMILY.
CC -1- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOG (CH) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL; AF237771; AAG37173.1; -
CC EMBL; AF325830; AAK49911.1; -
CC EMBL; AK001655; BAA91815.1; -
CC EMBL; BC016713; AAH16713.1; -
CC EMBL; BC014535; AAH14535.1; -
CC InterPro: IPR001715; Calponin_hom.
CC SMART: SM00033; CH; 2.
CC PROSITE: PS50021; CH; 2.
CC Cell adhesion; Cytoskeleton; Actin-binding; Repeat.
CC DOMAIN 95 201 CH 1.
CC FT DOMAIN 262 369 CH 2.
CC FT CONFLICT 11 11 V->A (IN REF. 4; AAH14535).
CC SEQUENCE 372 AA; 42243 MW; F48B5B183F8CEC CRC64;
OY 41 YNLFTK 46 10.3%; Score 6; DB 1; Length 372;
Db 362 YNLFTK 367 Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 15
PARA_MOUSE STANDARD: PRT; 372 AA.
AC Q9EPCL; Q9JG65;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Alpha-parvin (Actopaxin).
GN PARVA OR ACTP.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21102127; PubMed=11171322;
RA Olaki T.M., Noegel A.A., Korenbaum E.;
RT "Parvin, a 42 kDa focal adhesion protein, related to the alpha-actinin
RT superfamily";
RL J. Cell Sci. 114:525-538(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20576449; PubMed=11134073;
RA Nikolopoulos S.N., Turner C.E.;
RT "Actopaxin, a new focal adhesion protein that binds paxillin LD motifs
RT and actin and regulates cell adhesion";
RL J. Cell Biol. 151:1435-1448(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Osada N., Kusuda J., Tanuma R., Ito A., Hirata M., Sugano S.,

RA Hashimoto K.;
 RT "Isolation of full-length cDNA clones from mouse brain cDNA
 library";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Probably plays a role in the regulation of cell adhesion
 and cytoskeleton organization.
 CC -1- SUBUNIT: Interacts with integrin-linked protein kinase, actin and
 paxillin LDI and LD4 motifs.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; localized to focal adhesions.
 CC -1- SIMILARITY: BELONGS TO THE PARVIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOG (CH) DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF237774; AAC27175.1; -;
 DR EMBL; AF264766; AAC09803.1; -;
 DR EMBL; AB045321; BAA97981.1; -;
 DR MGD; MGI:1931144; Parva.
 DR InterPro: IPR001715; Calponin_hom.
 DR SMART; SM00033; CH; 2.
 DR PROSITE; PS50021; CH; 2.
 DR Cell adhesion; Cytoskeleton; Actin-binding; Repeat.
 FT DOMAIN 95 201 CH 1.
 FT DOMAIN 262 369 CH 2.
 FT CONFLICT 33 66 L -> H (IN REF. 3).
 FT CONFLICT 66 66 F -> S (IN REF. 3).
 SQ SEQUENCE 372 AA; 42329 MM; 1251F2586A1ACB64 CRC64;
 Query Match 10.3%; Score 6; DB 1; Length 372;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 41 YNLFTK 46
 |||||
 Db 362 YNLFTK 367
 RESULT 16
 PARA_RAT STANDARD; PRT; 372 AA.
 AC 09H97;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Alpha-parvin (Actopaxin).
 GN PARVA OR ACTP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20576449; PubMed=11134073;
 RA Nikolopoulos S.N., Turner C.E.;
 RT "Actopaxin, a new focal adhesion protein that binds paxillin LD motifs
 and actin and regulates cell adhesion";
 RL J. Cell Biol. 151:1435-1448(2000).
 CC -1- FUNCTION: Probably plays a role in the regulation of cell adhesion
 and cytoskeleton organization.
 CC -1- SUBUNIT: Interacts with integrin-linked protein kinase, actin and
 paxillin LDI and LD4 motifs.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; localized to focal adhesions.
 CC -1- SIMILARITY: BELONGS TO THE PARVIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOG (CH) DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF264765; AAC09802.2; -;
 DR InterPro: IPR001715; Calponin_hom.
 DR Pfam; PF00307; CH; 1.
 DR SMART; SM00033; CH; 2.
 DR PROSITE; PS50021; CH; 2.
 DR Cell adhesion; Cytoskeleton; Actin-binding; Repeat.
 FT DOMAIN 95 201 CH 1.
 FT DOMAIN 262 369 CH 2.
 SQ SEQUENCE 372 AA; 42291 MM; 7AAD24EBC25D094C CRC64;
 Query Match 10.3%; Score 6; DB 1; Length 372;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 41 YNLFTK 46
 |||||
 Db 362 YNLFTK 367
 RESULT 17
 PARA_RAT STANDARD; PRT; 378 AA.
 AC P54967;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Biotin synthase (EC 2.8.1.6) (Biotin synthetase).
 GN BIO2 OR BIOB OR AT2G43360 OR T01024.10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 GN NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=CV, LANDSBERG ERECTA;
 RX MEDLINE=96417082; PubMed=8819873;
 RA Weaver L.M., Yu F., Murtelle E.S., Nikolau B.J.;
 RT "Characterization of the cDNA and gene coding for the biotin synthase
 of Arabidopsis thaliana";
 RL Plant Physiol. 110:1021-1028(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=CV, COLUMBIA; TISSUE=Leaf;
 RX Patton D., Pacella M., Ward E.;
 RT Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=CV, COLUMBIA; TISSUE=Leaf;
 RX MEDLINE=96307524; PubMed=8680961;
 RA Balder P., Rufet M.L.;
 RT "Biotin synthesis in higher plants: isolation of a cDNA encoding
 Arabidopsis thaliana biob-gene product equivalent by functional
 complementation of a biotin auxotroph mutant bioB105 of Escherichia
 coli K12.";
 RL C. R. Acad. Sci., III, Sci. Vie 319:99-106(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX STRAIN=CV, COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Bueli C.R., Keichum K.A., Lee J.J., Rongning C.M., Koo H.L.,
 RA Moffat K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Unayam L.,
 RA Talon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
 RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
 RA Nerman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,

RA Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana";
RL Nature 402:761-768(1999).
CC -1- CATALYTIC ACTIVITY: Dethiobiotin + sulfur = biotin.
CC -1- PATHWAY: Biotin biosynthesis; last step.
CC -1- SIMILARITY: BELONGS TO THE BIOTIN AND LIPOIC ACID SYNTHETASES
FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: U24147; AAA80226.1; -;
CC EMBL: U31806; AAC49445.1; -;
CC EMBL: L34413; AAB39953.1; -;
CC EMBL: AC002335; AAB64312.1; -;
CC InterPro: IPR002684; Biotin_synth.
CC Pfam: PF01792; Biotin_synth; 1.
DR Biotin biosynthesis; Iron-sulfur; Transferase.
KW METAL 94 IRON-SULFUR (POTENTIAL).
FT METAL 98 IRON-SULFUR (POTENTIAL).
FT METAL 101 IRON-SULFUR (POTENTIAL).
SQ SEQUENCE 378 AA; 41681 MW; B102E477E7353762 CRC64;

Query Match 10.3%; Score 6; DB 1; Length 378;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 EYKEI 58
Db 158 EYKEI 163

RESULT 18
SPB7_HUMAN STANDARD; PRT; 380 AA.
AC 075635;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Megsin (TP55) (Serpin B7).
CN SERPINB7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9732616; PubMed=9182567;
RA Tsuchimoto M., Tsuruoka N., Ishida N., Kurihara T., Iwasa F.,
RA Yanashiro K., Rogi T., Kodama S., Katsuragi N., Adachi M.,
RA Katayama T., Nakao M., Yamauchi K., Hashino J., Hanyama M., Mura K.,
RA Nakanishi T., Nakazato H., Teramura M., Mizoguchi H., Yamaguchi N.;
RT "Purification, cDNA cloning, and characterization of a new serpin with
RT megakaryocyte maturation activity";
RL J. Biol. Chem. 272:15373-15380(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mesangial cells;
RX MEDLINE=98376492; PubMed=9710452;
RA Miyata T., Nangaku M., Suzuki D., Inagi R., Uragami K., Sakai H.,
RA Okubo K., Kurokawa K.;
RT "A mesangium-predominant gene, megin, is a new serpin upregulated in
RT IgA nephropathy";
RL J. Clin. Invest. 102:828-836(1998).
CC -1- FUNCTION: Might function as an inhibitor of Lys-specific
CC proteases. Might influence the maturation of megakaryocytes via
CC its action as a serpin.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- TISSUE SPECIFICITY: Predominantly expressed in mesangial cells.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: D88575; BAA31232.1; -;
CC EMBL: AF027866; AAC64506.1; -;
CC MIM: 603357; -;
CC HSSP: P05619; 1HE.
CC InterPro: IPR000215; Serpin.
DR Pfam: PF00079; Serpin; 1.
DR SMART: SM00093; SERPIN; 1.
DR PROSITE: PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor.
FT ACT_SITE 347 348 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 380 AA; 42904 MW; 9A2CDB6C63CFF605 CRC64;

Query Match 10.3%; Score 6; DB 1; Length 380;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 LYDAKV 34
Db 121 LYDAKV 126

RESULT 19
YA25_METJA STANDARD; PRT; 388 AA.
ID YA25_METJA
AC 058431;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ1025.
GN MJ1025.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RC MEDLINE=96337999; PubMed=6688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervlaeghe A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utermack T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii";
RL Science 273:1058-1073(1996).
RN [1]
RP -1- SIMILARITY: TO E. COLI YHAM.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: U67545; AAB99029.1; -;
CC TIGR: MJ1025; -;

KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 388 AA; 42644 MW; 80CA180E1061315C CRC64;

Query Match 10.3%; Score 6; DB 1; Length 388;
Best Local Similarity 100.0%; Pred. No. 31;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 DEIKGH 18
|||||
Db 247 DEIKGH 252

RESULT 20
YBIU_ECOLI STANDARD; PRT; 421 AA.

ID YBIU_ECOLI STANDARD; PRT; 421 AA.

AC P75791; 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae.

OX NCBI_TaxID=562;

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474(1997).

RM SEQUENCE FROM N.A.

RC STRAIN=K12;

RA MEDLINE=97061202; PubMed=8905232;

RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,

RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,

RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizouchi K.,

RA Mori H., Motomura K., Nakamura Y., Nishimoto Y., Saito N.,

RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,

RA Yano M., Horiuchi T.;

RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome

corresponding to the 12.7-28.0 min region on the linkage map.";

RL DNA Res. 3:137-155(1996).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC EMBL: AE000184; AAC73908.1; -

DR EMBL: D90719; BAA35502.1; -

DR EMBL: D90720; BAA35509.1; -

DR EcoGene: EG13326; ybiU.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 421 AA; 47329 MW; B3641337B6C62E48 CRC64;

Query Match 10.3%; Score 6; DB 1; Length 421;

Best Local Similarity 100.0%; Pred. No. 33;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 IKGHFP 20
|||||

Db 90 IKGHFP 95

RESULT 21
V218_FOWPV STANDARD; PRT; 461 AA.

ID V218_FOWPV STANDARD; PRT; 461 AA.

AC Q9517; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

OS Putative ankylin-repeat protein FPV218.

GN FPV218.

OC Fowlpox virus (FPV).

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Avipoxvirus.

OX NCBI_TaxID=10261;

RP SEQUENCE FROM N.A.

RC MEDLINE=20193820; PubMed=10729156;

RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;

RT "The genome of fowlpox virus.";

RL J. Virol. 74:3815-3831(2000).

CC -1 SIMILARITY: CONTAINS 12 ANK REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC EMBL: AF198100; AAF4562.1; -

DR InterPro: IPR002110; ANK.

DR Pfam: PF00023; ank; 8.

DR SMART: SM00248; ANK; 4.

DR PROSITE: PS50088; ANK_REPEAT; 3.

DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.

KW Hypothetical protein; Repeat; ANK repeat.

FT REPEAT 1 28 ANK 1.

FT REPEAT 31 61 ANK 2.

FT REPEAT 65 94 ANK 3.

FT REPEAT 96 116 ANK 4.

FT REPEAT 120 149 ANK 5.

FT REPEAT 153 182 ANK 6.

FT REPEAT 186 213 ANK 7.

FT REPEAT 217 248 ANK 8.

FT REPEAT 250 277 ANK 9.

FT REPEAT 281 312 ANK 10.

FT REPEAT 358 385 ANK 11.

FT REPEAT 431 460 ANK 12.

SQ SEQUENCE 461 AA; 52636 MW; E5892ABD916AB807 CRC64;

Query Match 10.3%; Score 6; DB 1; Length 461;

Best Local Similarity 100.0%; Pred. No. 36;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 50 YIIEYI 55
|||||

Db 438 YIIEYI 443

RESULT 22

SEC2_HELPY STANDARD; PRT; 503 AA.

AC O26074; 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

GN Protein-export membrane protein secD.

OS Helicobacter pylori (Campylobacter pylori).

OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;

OC Helicobacter.

OX NCBI_TaxID=210;

FN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-26695 / ATCC 700392;
 RX MEDLINE-9794467; PubMed-9252185;
 RA Tomb J.-F., White O., Kurlavsky A.R., Clayton R.A., Sutton G.G.,
 RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
 RA Nelson K., Quackenbush J., Zhou L., Kirness E.F., Peterson S.,
 RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
 RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
 RA Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,
 RA Cotton M.D., Weiman J.M., Fujit C., Bowman C., Matthey L., Wallin E.,
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.;
 RT "The complete genome sequence of the gastric pathogen *Helicobacter*
 RT *pylori*.";
 RL Nature 388:539-547(1997).
 CC -1- FUNCTION: INVOLVED IN PROTEIN EXPORT (BY SIMILARITY).
 CC -1- SUBUNIT: PART OF THE PROKARYOTIC PROTEIN TRANSLOCATION APPARATUS
 CC WHICH COMPRISE SECA, SECB, SECD, SECE, SECF, SECG AND SECH
 CC (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE SECD/SECF FAMILY. SECD FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AE000652; AAD08588.1; -
 CC TIGR: HP1550; -
 DR InterPro: IPR001036; ACR tran.
 DR InterPro: IPR003335; SECD_Secf.
 DR Pfam: PF02355; SECD_Secf; 1.
 DR PRINTS: PR00702; ACRIPLAVINRP.
 KW Protein transport; Translocation; Transmembrane; Membrane;
 KW Complete proteome.
 FT TRANSMEM 199 219 POTENTIAL.
 FT TRANSMEM 334 354 POTENTIAL.
 FT TRANSMEM 357 377 POTENTIAL.
 FT TRANSMEM 383 403 POTENTIAL.
 FT TRANSMEM 456 476 POTENTIAL.
 SQ SEQUENCE 503 AA; 54247 MW; 8541C291CA317086 CRC64;
 Query Match 10.3%; Score 6; DB 1; Length 503;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 31 DAKVY 36
 DB 226 DAKVY 231
 RESULT 23
 YAB5_METUA
 ID YAB5_METUA STANDARD: PRT: 509 AA.
 AC 058485; 199 219 POTENTIAL.
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein MJ1085.
 GN MJ1085.
 OS Methanococcus jannaschii.
 OC Archaeae; Euryarchaeota; Methanococcales; Methanococcaceae;
 CC Methanococcus.
 CC NCBI_TaxID=2190;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE-9633799; PubMed-8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,

RA Kurlavsky A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geophagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klein H.-P., Fraser C.M., Smith H.O., Weese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, *Methanococcus*
 RT *jannaschii*.";
 RL Science 273:1058-1073(1996).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U67551; AAB99095.1; -
 CC TIGR: MJ1085; -
 DR Hypothetical protein: Complete proteome.
 KW SEQUENCE 509 AA; 59471 MW; 8ACF7C4B0456049 CRC64;
 Query Match 10.3%; Score 6; DB 1; Length 509;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 51 IIEYK 56
 DB 435 IIEYK 440
 RESULT 24
 GUX1_TRIVI
 ID GUX1_TRIVI STANDARD: PRT: 513 AA.
 AC P19355;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Exoglucanase I precursor (EC 3.2.1.91) (Exocelllobiohydrolase) (1,4-
 DE beta-cellobiohydrolase).
 GN CBHL.
 OS Trichoderma viride.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreales; mitosporic Hypocreales; Trichoderma.
 CC NCBI_TaxID=5347;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE-91016856; PubMed-2216737;
 RA Cheng C., Tsukagoshi N., Ueda S.;
 RT "Nucleotide sequence of the cellobiohydrolase gene from *Trichoderma*
 RT *viride*.";
 RL Nucleic Acids Res. 18:5559-5559(1990).
 CC -1- FUNCTION: THE BIOLOGICAL CONVERSION OF CELLULOSE TO GLUCOSE
 CC GENERALLY REQUIRES THREE TYPES OF HYDROLYTIC ENZYMES:
 CC (1) ENDOGLYCANASES WHICH CUT INTERNAL BETA-1,4-GLUCOSIDIC BONDS;
 CC (2) EXOCELLULOBIODIOLASES THAT CUT THE DISACCHARIDE CELLOBIOSE
 CC FROM THE NONREDUCING END OF THE CELLULOSE POLYMER CHAIN;
 CC (3) BETA-1,4-GLUCOSIDASES WHICH HYDROLYZE THE CELLOBIOSE AND OTHER
 CC SHORT CELLO-OLIGOSACCHARIDES TO GLUCOSE.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glucosidic linkages
 CC in cellulose and cellobiose, releasing cellobiose from the non-
 CC reducing ends of the chains.
 CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY C (FAMILY 7 OF GLYCOSYL
 CC HYDROLASES).
 CC -1- SIMILARITY: CONTAINS 1 FUNGAL-TYPE CELLULOSE-BINDING DOMAIN (CBD).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

FT TRANSMEM 478 498 POTENTIAL.
SQ SEQUENCE 526 AA; 56796 MM; 5BAC9E05794782A CRC64;
Query Match 10.3%; Score 6; DB 1; Length 526;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 31 DAKVY 36
|||||
DB 248 DAKVY 253

RESULT 27
HG11_CANAL STANDARD; PRT; 545 AA.
AC 074713;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE High-affinity glucose transporter.
GN Hgt1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10261;
RA Varma A., Singh B.B., Karnani N., Lichtenberg H., Magee B.B.,
RA Hofer M., Prasad R.;
RT "Molecular characterization of glucose transporter, CAHG1, of a
pathogenic yeast Candida albicans."
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: HIGH-AFFINITY GLUCOSE TRANSPORTER.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y16834; CAAT6406.1; -;
DR InterPro; IPR003663; Sugar transporter.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTNSPORT.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Transport; Glycoprotein.
FT DOMAIN 1 28
FT TRANSMEM 29 49
FT DOMAIN 50 71
FT TRANSMEM 72 92
FT DOMAIN 93 99
FT TRANSMEM 100 120
FT TRANSMEM 121 124
FT TRANSMEM 125 145
FT TRANSMEM 146 156
FT TRANSMEM 157 177
FT TRANSMEM 178 191
FT TRANSMEM 192 212
FT TRANSMEM 213 290
FT TRANSMEM 291 311
FT TRANSMEM 312 316
FT TRANSMEM 317 337
FT TRANSMEM 338 344
FT TRANSMEM 345 365
FT TRANSMEM 366 394
FT TRANSMEM 395 415

FT DOMAIN 416 432 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 433 453 11 (POTENTIAL).
FT DOMAIN 454 459 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 460 480 12 (POTENTIAL).
FT DOMAIN 481 545 CYTOPLASMIC (POTENTIAL).
FT CARBOHD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 545 AA; 60670 MM; A9A0FD77DC07080E CRC64;
Query Match 10.3%; Score 6; DB 1; Length 545;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 43 LFTKY 48
|||||
DB 270 LFTKY 275

RESULT 28
SYR_CORGL STANDARD; PRT; 550 AA.
AC P35668; P41253;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Arginyl-tRNA synthetase (EC 6.1.1.19) (Arginine--tRNA ligase) (ARGRS).
GN ARGS.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;
OC Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A. / AS019;
RC STRAIN=ATCC 13059 /
RA MEDLINE=91186817; PubMed=2082143;
RT Marcel T., Archer J.A.C., Mengin-Lecreux D., Sinskey A.J.;
RT "Nucleotide sequence and organization of the upstream region of the
Corynebacterium glutamicum lysA gene."
RL Mol. Microbiol. 4:1819-1830(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13869;
RA MEDLINE=94042911; PubMed=8226683;
RT Ogutza J.A., Malumbres M., Eriani G., Pisabarro A., Mateos L.M.,
RA Martin F., Martin J.F.;
RT "A gene encoding arginyl-tRNA synthetase is located in the upstream
region of the lysA gene in Brevibacterium lactofermentum: regulation
of argS-lysA cluster expression by arginine."
RL J. Bacteriol. 175:7356-7362(1993).
RN [3]
RP IDENTIFICATION.
RX MEDLINE=93268096; PubMed=8497194;
RT Sharp P.M., Mitchell K.J.;
RT "Corynebacterium glutamicum arginyl-tRNA synthetase."
RL Mol. Microbiol. 8:200-206(1993).
CC -1- CATALYTIC ACTIVITY: ATP + L-arginine + tRNA(Arg) = AMP +
diphosphate + L-arginyl-tRNA(Arg).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X54740; CA38537.1; ALT_INIT.
DR EMBL; Z21501; CAAT9710.1; -;
DR PIR; S12227; S12227.

DR PIR: S42850; S42850.
DR InterPro: IPR001278; tRNA-synt_1d.
DR InterPro: IPR001412; tRNA-synt_1.
DR Pfam: PF00750; tRNA-synt_1d: 1.
DR PRINTS: PR01038; TRNASYNTHARG.
DR PROSITE: PS00178; AA-TRNA-LIGASE_I; 1.
KW Anticodon-1 tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 132 140 "HIGH" REGION.
FT SITE 374 378 "KMSKS" REGION.
FT BINDING 377 377 ATP (BY SIMILARITY).
FT CONFLICT 355 355 G -> D (IN REF. 2).
FT CONFLICT 412 412 I -> M (IN REF. 2).
FT CONFLICT 513 513 V -> A (IN REF. 2).
FT CONFLICT 540 540 H -> R (IN REF. 2).
SQ SEQUENCE 550 AA; 59723 MW; 3AF724BDE8DC4C1 CRC64;
Query Match 10.3%; Score 6; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 53 EYKEI 58
DB 198 EYKEI 203
RESULT 29
YAY3_SCHPO
ID YAY3_SCHPO STANDARD; PRT; 649 AA.
AC 010211:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHETICAL 74.5 kDa protein C4H3.03C in chromosome I.
GN SPAC4H3.03C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Murphy L., Harris D., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: Z69380; CA93342.1; -
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 265 285 POTENTIAL.
FT TRANSMEM 584 584 POTENTIAL.
FT TRANSMEM 626 646 POTENTIAL.
SQ SEQUENCE 649 AA; 74488 MW; 0C97C10E603EB33D CRC64;
Query Match 10.3%; Score 6; DB 1; Length 649;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 24 SMFEDL 29
DB 594 SMFEDL 599
RESULT 30
ALT_BPT4 STANDARD; PRT; 682 AA.
ID ALT_BPT4

AC P12726;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE MAD--protein ADP-ribosyltransferase (EC 2.4.2.-) (Alt protein).
GN ALT.
OS Bacteriophage T4.
OS Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like phages.
OX NCBI_TaxID=10665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C;
KW MEDLINE=89386005; PubMed=2506526;
RA Hilse D., Koch T., Rueger W.;
RL "Nucleotide sequence of the alt gene of bacteriophage T4.";
RN Nucleic Acids Res. 17:6731-6731(1989).
RP REVISIONS.
RX MEDLINE=94330139; PubMed=8053153;
RA Koch T., Rueger W.;
RL "The ADP-ribosyltransferases (gpAlt) of bacteriophages T2, T4, and T6:
RL sequencing of the genes and comparison of their products.";
RN Virology 203:294-298(1994).
RN [3]
RP SEQUENCE FROM N.A.
RA Kutter E., Arisaka F., Kunisawa T., Tsugita A., Mosig G.,
RA Mesyanzhinov V., Ruger W., Stidham T., Thomas E.;
RL "Bacteriophage T4 genome analysis.";
CC Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMPONENT OF THE PHAGE HEAD: ADP-RIBOSYLATES ONE OF THE
CC 2 ALPHA-SUBUNITS OF THE BACTERIAL RNA POLYMERASE.
CC -!- SUBCELLULAR LOCATION: THIS PROTEIN IS INJECTED FROM THE VIRION
CC INTO THE BACTERIAL CELL.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X15811; CA933807.1; -
DR EMBL: AF158101; AAD42533.1; -
DR PIR: J00096; SXBPT4.
KW Transferase; Glycosyltransferase.
RP PROPEP 1 6
FT CHAIN 7 682
FT CONFLICT 22 38
FT CONFLICT 266 266
FT CONFLICT 527 527
FT CONFLICT 636 636
FT CONFLICT 658 661
FT CONFLICT 665 669
SQ SEQUENCE 682 AA; 75817 MW; DE31BDC56AEC427 CRC64;
Query Match 10.3%; Score 6; DB 1; Length 682;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 KKYAYI 51
DB 153 KKYAYI 158
RESULT 31
ALT_BPT2 STANDARD; PRT; 698 AA.
ID ALT_BPT2
AC Q38424;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE NAD--protein ADP-ribosyltransferase (EC 2.4.2.-) (Alt protein).
GN ALT
OS Bacteriophage T2.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OX T4-like phages.
NCBI_TaxID=10664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WT.
RX MEDLINE=94330139; PubMed=8053153;
RA Koch T., Rueger W.;
RT "The ADP-ribosyltransferases (gpAlt) of bacteriophages T2, T4, and
T6: sequencing of the genes and comparison of their products.";
RL Virology 203:294-298(1994).
CC -1- FUNCTION: COMPONENT OF THE PHAGE HEAD; ADP-RIBOSYLATES ONE OF THE
2 ALPHA-SUBUNITS OF THE BACTERIAL RNA POLYMERASE.
CC -1- SUBCELLULAR LOCATION: THIS PROTEIN IS INJECTED FROM THE VIRION
CC INTO THE BACTERIAL CELL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X69893; CAA49517.1; -
RW Transferase; Glycosyltransferase.
SQ SEQUENCE 698 AA; 77996 MW; B2D0BAEB729457C6 CRC64;
Query Match 10.3%; Score 6; DB 1; Length 698;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 KKYAYI 51
DB 154 KKYAYI 159
RESULT 32
ID ALT_BPT6 STANDARD; PRT; 698 AA.
AC Q38433;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE NAD--protein ADP-ribosyltransferase (EC 2.4.2.-) (Alt protein).
GN ALT
OS Bacteriophage T6.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like phages.
OX NCBI_TaxID=10666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WT.
RX MEDLINE=94330139; PubMed=8053153;
RA Koch T., Rueger W.;
RT "The ADP-ribosyltransferases (gpAlt) of bacteriophages T2, T4, and
T6: sequencing of the genes and comparison of their products.";
RL Virology 203:294-298(1994).
CC -1- FUNCTION: COMPONENT OF THE PHAGE HEAD; ADP-RIBOSYLATES ONE OF THE
2 ALPHA-SUBUNITS OF THE BACTERIAL RNA POLYMERASE.
CC -1- SUBCELLULAR LOCATION: THIS PROTEIN IS INJECTED FROM THE VIRION
CC INTO THE BACTERIAL CELL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X69894; CAA49518.1; -
RW Transferase; Glycosyltransferase.
SQ SEQUENCE 698 AA; 77947 MW; 1704DF87A75CD835 CRC64;
Query Match 10.3%; Score 6; DB 1; Length 698;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 KKYAYI 51
DB 154 KKYAYI 159
RESULT 33
ID TE80_TETTH STANDARD; PRT; 719 AA.
AC Q94818;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Telomerase component p80 (EC 2.7.7.-).
OS Tetrahymena thermophila.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC Tetrahymenina; Tetrahymena.
OX NCBI_TaxID=5911;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=95292335; PubMed=7774009;
RA Collins K., Kobayashi R., Greider C.W.;
RT "Purification of Tetrahymena telomerase and cloning of genes encoding
the two protein components of the enzyme.";
RL Cell 81:677-686(1995).
CC -1- FUNCTION: RIBONUCLEOPROTEIN DNA POLYMERASE THAT CATALYZES THE DE
NOVO SYNTHESIS OF TELOMERIC SIMPLE SEQUENCE REPEATS. P80 BINDS
TIGHTLY AND SPECIFICALLY TO THE TELOMERASE RNA SUGGESTING ITS
ASSOCIATION WITH A REGION OF RNA SECONDARY STRUCTURE.
CC -1- SUBUNIT: TELOMERASE CONSIST OF TWO SUBUNIT, P80 AND P95 THAT FORM
A 1:1:1 COMPLEX WITH THE 159 NT TELOMERASE RNA.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U25641; AAC46601.1; -
RW Transferase; RNA-directed DNA polymerase; Telomere; Nuclear protein;
KW DNA-binding.
SQ SEQUENCE 719 AA; 82351 MW; 8A945A71189CA99C CRC64;
Query Match 10.3%; Score 6; DB 1; Length 719;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 DEIKGH 18
DB 504 DEIKGH 509
RESULT 34
ID CAD6_RAT STANDARD; PRT; 789 AA.
AC P55280;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cadherin-6 precursor (Kidney-cadherin) (K-cadherin).
GN CDH6 OR KCAD.

OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ACI; TISSUE=Kidney;
 RX MEDLINE=94243827; Pubmed=8187093;
 RA Xiang Y.Y., Tanaka M., Suzuki M., Igarashi H., Kiyokawa E., Naito Y.,
 RA Ohtawara Y., Shen Q., Sugimura H., Kino I.;
 RT "Isolation of complementary DNA encoding K-cadherin, a novel rat
 RT cadherin preferentially expressed in fetal kidney and kidney
 RT carcinoma.";
 RL Cancer Res. 54:3034-3041(1994).
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN KIDNEY AND BRAIN.
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D25290; BAA04975.1; -;
 CC HSSP: P15116; INCU.
 DR InterPro: IPR002126; Cadherin.
 DR InterPro: IPR000233; Cadherin_C-term.
 DR Pfam: PF00028; cadherin_5.
 DR Pfam: PF01049; Cadherin_C-term; 1.
 DR PRINTS: PR00205; CADHERIN.
 DR SMART: SM00112; CA; 5.
 DR PROSITE: PS00232; CADHERIN_1; 3.
 DR PROSITE: PS50268; CADHERIN_2; 5.
 DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KW Signal.
 FT SIGNAL. 1 18
 FT PROPEP 19 53
 FT CHAIN 54 789
 FT DOMAIN 54 615
 FT TRANSMEM 616 636
 FT DOMAIN 637 789
 FT DOMAIN 54 159
 FT DOMAIN 160 268
 FT DOMAIN 269 383
 FT DOMAIN 384 487
 FT DOMAIN 487 608
 FT DOMAIN 608 636
 FT CARBOHYD 399 399
 FT CARBOHYD 437 437
 FT CARBOHYD 455 455
 FT CARBOHYD 536 536
 FT SEQUENCE 789 AA; 88340 MW; E27474F0B07FE403 CRC64;
 Query Match 10.38; Score 6; DB 1; Length 789;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 32 AKVYYS 37
 DB 195 AKVYYS 200
 RESULT 35
 CAD9_HUMAN STANDARD; PRT; 789 AA.
 AC Q9ULB4;

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cadherin-9 precursor.
 GN CDH9.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=10861224;
 RA Shimoyama Y., Tsujimoto G., Kitajima M., Natori M.;
 RT "Identification of three human type-II classic cadherins and frequent
 RT heterophilic interactions between different subclasses of type-II
 RT classic cadherins.";
 RL Blochm. J. 349:159-167(2000).
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB035302; BAA87416.1; -;
 CC HSSP: P15116; INCU.
 DR InterPro: IPR002126; Cadherin.
 DR InterPro: IPR000233; Cadherin_C-term.
 DR Pfam: PF00028; cadherin_5.
 DR Pfam: PF01049; Cadherin_C-term; 1.
 DR PRINTS: PR00205; CADHERIN.
 DR SMART: SM00112; CA; 5.
 DR PROSITE: PS00232; CADHERIN_1; 2.
 DR PROSITE: PS50268; CADHERIN_2; 5.
 DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KW Signal.
 FT SIGNAL. 1 21
 FT PROPEP 22 53
 FT CHAIN 54 789
 FT DOMAIN 54 615
 FT TRANSMEM 616 636
 FT DOMAIN 637 789
 FT DOMAIN 54 159
 FT DOMAIN 160 268
 FT DOMAIN 269 383
 FT DOMAIN 384 486
 FT DOMAIN 487 608
 FT CARBOHYD 255 255
 FT CARBOHYD 437 437
 FT CARBOHYD 455 455
 FT CARBOHYD 536 536
 FT SEQUENCE 789 AA; 859853275644344F CRC64;
 Query Match 10.38; Score 6; DB 1; Length 789;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 32 AKVYYS 37
 DB 195 AKVYYS 200
 RESULT 36
 CAD6_CHICK STANDARD; PRT; 790 AA.
 AC CAD6_CHICK

AC 090762;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cadherin-6 precursor (Cadherin-6b) (c-cad6b).
 GN CDH6.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 CC NCBI_TaxID=9031;
 RX MEDLINE=95309115; PubMed=7540531;
 RA Nakagawa S., Takeichi M.;
 RT "Neural crest cell-cell adhesion controlled by sequential and
 RT subpopulation-specific expression of novel cadherins.";
 RL Development 121:1321-1332(1995).
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -1- DEVELOPMENTAL STAGE: FIRST EXPRESSED IN SPLANCHNIC MESODERM OF
 CC STAGE 4 EMBRYOS. AT STAGE 6, STRONGLY EXPRESSED ALONG THE NEURAL
 CC FOLD IN A REGION CORRESPONDING TO THE FUTURE NEURAL CREST.
 CC EXPRESSION IN THE NEURAL FOLD CONTINUES DURING CLOSURE OF THE
 CC NEURAL TUBE BUT DIMINISHES AFTER NEURAL CREST CELLS HAVE LEFT THE
 CC NEURAL TUBE.
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D42149; BAA07720.1; -.
 DR HSP: P15116; INCU.
 DR InterPro: IPR002126; Cadherin.
 DR InterPro: IPR000233; Cadherin_C-term.
 DR Pfam: PF00028; cadherin_5.
 DR Pfam: PF01049; cadherin_C-term; 1.
 DR PRINTS: PR00205; CADHERIN.
 DR SMART: SM00112; CA: 5.
 DR PROSITE: PS00232; CADHERIN_1; 3.
 DR PROSITE: PS50268; CADHERIN_2; 5.
 KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KW Signal.
 FT SIGNAL 1 30
 FT PROPEP 31 53 POTENTIAL.
 FT CHAIN 54 790 POTENTIAL.
 FT DOMAIN 54 615 CADHERIN-6.
 FT TRANSEM 616 636 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 637 790 POTENTIAL.
 FT DOMAIN 54 159 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 54 159 CADHERIN 1.
 FT DOMAIN 160 268 CADHERIN 2.
 FT DOMAIN 269 383 CADHERIN 3.
 FT DOMAIN 384 486 CADHERIN 4.
 FT DOMAIN 487 608 CADHERIN 5.
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 455 455 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 790 AA; 88659 MW; 0FD3756749DB5C5C5 CRC64;

Query Match 10.38; Score 6; DB 1; Length 790;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

32 AKVYS 37
 11111
 195 AKVYS 200

RESULT 37
 CAD6_HUMAN
 ID CAD6_HUMAN STANDARD: PRT: 790 AA.
 AC P55285; O9BMS0;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cadherin-6 precursor (Kidney-cadherin) (K-cadherin).
 GN CDH6.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 CC NCBI_TaxID=9606;
 RX MEDLINE=95262134; PubMed=7743525;
 RA Shimoyama Y., Gotoh M., Terasaki T., Kitajima M., Hirohashi S.;
 RT "Isolation and sequence analysis of human cadherin-6 complementary
 RT DNA for the full coding sequence and its expression in human
 RT carcinoma cells.";
 RL Cancer Res. 55:2206-2211(1995).
 RN [1]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RP MEDLINE=95262134; PubMed=7743525;
 RA Shimoyama Y., Gotoh M., Terasaki T., Kitajima M., Hirohashi S.;
 RT "Isolation and sequence analysis of human cadherin-6 complementary
 RT DNA for the full coding sequence and its expression in human
 RT carcinoma cells.";
 RL Cancer Res. 55:2206-2211(1995).
 RN [2]
 RN SEQUENCE FROM N.A. (ISOFORM 2).
 RP TISSUE=Kidney;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE OF 377-790 FROM N.A. (ISOFORM 1).
 RP TISSUE=Fetal brain;
 RX MEDLINE=91283540; PubMed=2059658;
 RA Suzuki S., Sano K., Tanihara H.;
 RT "Diversity of the cadherin family: evidence for eight new cadherins
 RT in nervous tissue.";
 RL Cell Regul. 2:261-270(1991).
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN, CEREBELLUM, AND.
 CC KIDNEY, LUNG, PANCREAS, AND GASTRIC MUCOSA SHOW A WEAK EXPRESSION.
 CC ALSO EXPRESSED IN CERTAIN LIVER AND KIDNEY CARCINOMAS.
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D31784; BAA06562.1; -.
 DR EMBL: BC000019; AAH00019.1; -.
 DR HSP: P15116; INCU.
 DR MIM: 603007; -.
 DR InterPro: IPR002126; Cadherin.
 DR InterPro: IPR000233; Cadherin_C-term.
 DR Pfam: PF00028; cadherin_5.
 DR Pfam: PF01049; cadherin_C-term; 1.
 DR PRINTS: PR00205; CADHERIN.
 DR SMART: SM00112; CA: 5.
 DR PROSITE: PS00232; CADHERIN_1; 3.
 DR PROSITE: PS50268; CADHERIN_2; 5.
 KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;

KW Signal; Alternative splicing.
 FT SIGNAL 1 18
 FT PROPEP 19 53
 FT CHAIN 54 790
 FT DOMAIN 54 615
 FT TRANSMEM 616 636
 FT DOMAIN 637 790
 FT DOMAIN 54 159
 FT DOMAIN 160 268
 FT DOMAIN 269 383
 FT DOMAIN 384 486
 FT DOMAIN 487 608
 FT CARBOHYD 49 255
 FT CARBOHYD 255 399
 FT CARBOHYD 399 437
 FT CARBOHYD 437 455
 FT CARBOHYD 455 536
 FT CARBOHYD 536 663
 FT VARSPLIC 628 663
 FT VARSPLIC 664 790
 FT CONFLICT 421 421
 FT CONFLICT 425 425
 FT SEQUENCE 790 AA; 88308 MW; C175004FC8A61100 CRC64;
 Query Match 10.3%; Score 6; DB 1; Length 790;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 32 AKVYVS 37
 Db 195 AKVYVS 200
 RESULT 38
 CAD6_MOUSE STANDARD; PRT; 790 AA.
 ID CAD6_MOUSE P70393;
 AC P97326; (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cadherin-6 precursor (Kidney-cadherin) (K-cadherin).
 GN CDH6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9271342; PubMed=9126293;
 RA Inoue T., Chisaka O., Matsunami H., Takeichi M.;
 RT "Cadherin-6 expression transiently delineates specific rhombomeres,
 RT other neural tube subdivisions, and neural crest subpopulations in
 RT mouse embryos.";
 RL Dev. Biol. 183:183-194(1997).
 RN [2]
 RP SEQUENCE OF 479-666 FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Kidney;
 RA Faulkner-Jones B.E., Dziadek M.A.;
 RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP DEVELOPMENTAL STAGE.
 RC STRAIN=C57BL/6; TISSUE=Testis;
 RX MEDLINE=9703837; PubMed=8879495;
 RA Munro S.B., Blaschuk O.W.;
 RT "A comprehensive survey of the cadherins expressed in the testes of
 RT fetal, immature, and adult mice utilizing the polymerase chain
 RT reaction.";
 RL Biol. Reprod. 55:822-827(1996).
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE

CC SORTING OF HETEROGENEOUS CELL TYPES.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN FETAL, NEWBORN AND 7-DAY-OLD
 CC TESTIS BUT NOT IN 21-DAY-OLD OR ADULT TESTIS.
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D82029; BA01516.1; -;
 CC EMBL: U67399; AAB07550.1; -;
 CC HSSP: P15116; INCI;
 CC MGD: MG1107435; Cdh6.
 CC InterPro: IPR002126; Cadherin.
 CC InterPro: IPR000233; Cadherin_C-term.
 CC Pfam: PF00028; cadherin; 5.
 CC ProSite: PS50268; Cadherin_2; 5.
 CC PRINTS: PR00205; CADHERIN.
 CC SMART: SM00112; CA; 5.
 CC PROSITE: PS00232; CADHERIN_1; 3.
 CC PROSITE: PS50268; CADHERIN_2; 5.
 CC Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KW Signal.
 FT SIGNAL 1 18
 FT PROPEP 19 53
 FT CHAIN 54 790
 FT DOMAIN 54 615
 FT TRANSMEM 616 636
 FT DOMAIN 637 790
 FT DOMAIN 54 159
 FT DOMAIN 160 268
 FT DOMAIN 269 383
 FT DOMAIN 384 486
 FT DOMAIN 487 608
 FT CARBOHYD 255 399
 FT CARBOHYD 399 437
 FT CARBOHYD 437 455
 FT CARBOHYD 455 536
 FT CARBOHYD 536 534
 FT CONFLICT 534 536
 FT CONFLICT 536 598
 FT CONFLICT 598 613
 FT CONFLICT 613 640
 FT SEQUENCE 790 AA; 88374 MW; FE3158AF165C579 CRC64;
 Query Match 10.3%; Score 6; DB 1; Length 790;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 32 AKVYVS 37
 Db 195 AKVYVS 200
 RESULT 39
 CHS1_PHYBL STANDARD; PRT; 841 AA.
 ID CHS1_PHYBL P87073;
 AC P87073;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Chitin synthase 1 (EC 2.4.1.16) (Chitin-UDP acetylglucosaminyl
 DE transferase 1) (Class-II chitin synthase 1).
 GN CHS1.
 OS Phycomyces blakesleeanus.
 OC Eukaryota; Fungi; Zygomycetes; Mucorales; Mucoraceae;
 CC Phycomyces.


```

OX NCBI_TaxID=4837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 1555;
RA Miyazaki A., Ootaki T.;
RN Submitted (May-1997) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE OF 172-370 FROM N.A.
RC STRAIN=NRRL 1555;
RA Miyazaki A., Momany M., Szaniszló P.J., Jayaram M., Ootaki T.;
RT "Chitin synthase-encoding gene(s) of the Zygomycete fungus Phycomyces blakesleeanus.";
RL Gene 134:129-134(1993).
CC -1- FUNCTION: PLAYS A MAJOR ROLE IN CELL WALL BIOGENESIS.
CC -1- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + [(1,4)-(N-acetyl-beta-D-glucosaminyl)](N+1).
CC -1- SUBCELLULAR LOCATION: Plasma membrane-bound.
CC -1- SIMILARITY: BELONGS TO THE CHITIN SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB003043; BAA19857.1; -
DR InterPro: IPR002923; Chitin_synth.
DR Pfam: PF01644; Chitin_synth. 1.
DR ProDom: PD002998; Chitin_synth. 1.
KW Transferrase; Glycosyltransferase; Transmembrane; Cell wall;
KW Multigene family.
FT TRANSMEM 302 322 POTENTIAL.
FT TRANSMEM 385 405 POTENTIAL.
FT TRANSMEM 526 546 POTENTIAL.
FT TRANSMEM 564 584 POTENTIAL.
FT TRANSMEM 602 622 POTENTIAL.
FT TRANSMEM 644 664 POTENTIAL.
FT TRANSMEM 673 693 POTENTIAL.
FT TRANSMEM 778 798 POTENTIAL.
FT TRANSMEM 816 836 POTENTIAL.
FT TRANSMEM 179 179 E -> EDE (IN REF. 2).
FT TRANSMEM 199 200 RV -> HI (IN REF. 2).
FT TRANSMEM 352 352 A -> G (IN REF. 2).
FT TRANSMEM 841 AA; 95226 MW; 71CD6C09ACB68B CRC64;
SQ SEQUENCE

Query Match 10.3%; Score 6; DB 1; Length 841;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 MFEEDLY 30
Db 516 MFEEDLY 521

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=90269039; PubMed=2161319;
RA Choe M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA Horsnell T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,
RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Bartell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169.";
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -1- FUNCTION: THIS PROTEIN MAY AFFECT TRANSDUCTION OF THE VIRUS
CC GLYCOPROTEINS TO MEMBRANES. IT IS INVOLVED IN CAPSID MATURATION.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES PTPP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17403; CA35371.1; -
DR PIR; S09819; WMBE56.
DR InterPro: IPR000501; Proc_transport.
DR Pfam; PF01366; PTPP; 1.
KW Capsid assembly.
SQ SEQUENCE 850 AA; 95868 MW; C32A91906DAFEE7D CRC64;

Query Match 10.3%; Score 6; DB 1; Length 850;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 EYNLEF 45
Db 280 EYNLEF 285

RESULT 41
CHSL_RHTOL
ID CHSL_RHTOL STANDARD; PRT; 858 AA.
AC P30594;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Chitin synthase 1 (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl
DE transferase 1).
DE CHSL.
GN Rhizopus oligosporus.
OS Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC Rhizopus.
OX NCBI_TaxID=4847;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95036875; PubMed=7765484;
RA Motoyama T., Sudoh M., Horiuchi H., Ohta A., Takagi M.;
RT "Isolation and characterization of two chitin synthase genes of
RT Rhizopus oligosporus.";
RL Biosci. Biotechnol. Biochem. 58:1685-1693(1994).
CC -1- FUNCTION: PLAYS A MAJOR ROLE IN CELL WALL BIOGENESIS.
CC -1- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + [(1,4)-(N-acetyl-beta-D-glucosaminyl)](N+1).
CC -1- SUBCELLULAR LOCATION: Plasma membrane-bound.
CC -1- SIMILARITY: BELONGS TO THE CHITIN SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10159; BAA1023.1; -

```

```

DR PIR: JC2308; JC2308.
DR InterPro: IPR002923; Chitin_synth.
DR PROSITE: PS001173; Glycosyl_transf_2.
DR Pfam: PF01644; Chitin_synth; 1.
DR ProDom: PD002998; Chitin_synth; 1.
DR Transfaser: Glycosyltransferase; Transmembrane, Cell wall;
KW Multigene family.
SQ SEQUENCE 858 AA; 97057 MW; EE1E6197F0DE70B9 CRC64;

Query Match 10.3%; Score 6; DB 1; Length 858;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 MFEDLY 30
Db 530 MFEDLY 535

RESULT 42
SYL_THEME STANDARD; PRT; 919 AA.
ID SYL_THEME
AC P46213;
DT 01-NOV-1995 (Rel. 32, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Isolation of a trna synthetase (EC 6.1.1.5) (Isolation of a trna synthetase)
DE (11ers).
GN ILES OR TMI361.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_Taxid=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RA MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima."
RT Nature 399:323-329(1999).
RN [2]
RP SEQUENCE OF 95-599 FROM N.A.
RC MEDLINE=95223956; PubMed=7708661;
RA Brown J.R., Doolittle W.F.;
RT "Root of the universal tree of life based on ancient aminoacyl-tRNA
RT synthetase gene duplications."
RT Proc. Natl. Acad. Sci. U.S.A. 92:2441-2445(1995).
CC -1- CATALYTIC ACTIVITY: ATP + L-isoleucine + tRNA(Ile) = AMP +
CC di-phosphate + L-isoleucyl-tRNA(Ile).
CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001190; AAD36431.1; -.
CC EMBL: L37104; AAC41448.1; -.
CC HSSP: PA1972; 1FFY.
CC TIGR: TMI361; -.
CC InterPro: IPR002300; tRNA-synt_1a.
CC InterPro: IPR001412; tRNA-synt_1i.
CC InterPro: IPR002301; tRNA-synt_1le.

```

```

DR Pfam: PF00133; tRNA-synt_1; 1.
DR PRINTS: PR00984; TRNASYNTHILE.
DR PROSITE: PS00178; AA TRNA LIGASE I; 1.
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Metal-binding; Zinc; Complete proteome.
DR SITE 57 67 "HIGH" REGION.
DR SITE 57 67 "KMSKS" REGION.
DR BINDING 594 598 ATP (BY SIMILARITY).
DR BINDING 597 597
SQ SEQUENCE 919 AA; 107155 MW; 40E4D0876010C385 CRC64;

Query Match 10.3%; Score 6; DB 1; Length 919;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKRYTV 8
Db 775 RKRYTV 780

RESULT 43
AMPE_MOUSE STANDARD; PRT; 945 AA.
ID AMPE_MOUSE
AC P16406;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutamyl aminopeptidase (EC 3.4.11.7) (EAP) (Aminopeptidase A) (APA)
DE (BP-1/6C3 antigen).
GN ENPEP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=90139003; PubMed=1689065;
RA Wu Q., Lahli J.M., Air G.M., Burrows P.D., Cooper M.D.;
RA "Molecular cloning of the murine BP-1/6C3 antigen: a member of the
RA zinc-dependent metalloproteinase family."
RA Proc. Natl. Acad. Sci. U.S.A. 87:993-997(1990).
RT -1- FUNCTION: APPEARS TO HAVE A ROLE IN THE CATABOLIC PATHWAY OF THE
RT RENIN-ANGIOTENSIN SYSTEM. PROBABLY PLAYS A ROLE IN REGULATING
RT GROWTH AND DIFFERENTIATION OF EARLY B-LINEAGE CELLS.
RT -1- CATALYTIC ACTIVITY: Release of a N-terminal glutamate (and to a
RT lesser extent aspartate) from a peptide.
CC -1- COFACTOR: REPORTEDLY ZINC-INDEPENDENT, REQUIRES CALCIUM FOR FULL
CC ACTIVITY.
CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: TYPE II membrane protein.
CC -1- TISSUE SPECIFICITY: EARLY B-LINEAGE CELLS AND CERTAIN STROMAL CELL
CC OF HEMOPOIETIC TISSUES. ALSO EXPRESSED BY CAPILLARY ENDOTHELIAL
CC CELLS, PLACENTA, AND EPITHELIAL CELLS OF THE INTESTINE AND
CC PROXIMAL RENAL TUBULES.
CC -1- PTM: PHOSPHORYLATED.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M29961; ABA47732.1; -.
CC PIR: S30398; S30398.
CC MEROPS: M01.003; -.
CC MGD: MGI:106645; Enpep.
CC InterPro: IPR001930; Aladipeptase.
CC InterPro: IPR000130; Zn_MTPeptide.
CC Pfam: PF01433; Peptidase_M1; 1.
CC PRINTS: PR00756; ALADIPITASE.
CC PROSITE: PS00142; ZINC_PROTEASE; 1.

```

KW Hydrolase; Metalloprotease; Aminopeptidase; Zinc; glycoprotein;
 KW B-cell; Antigen; Transmembrane; Phosphorylation; Signal-anchor;
 KW Calcium.
 FT DOMAIN 1 17 CYTOPLASMIC (POTENTIAL)
 FT TRANSSEM 18 40 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT
 FT DOMAIN 41 945
 FT METAL 385 385 EXTRACELLULAR (POTENTIAL).
 FT ACT_SITE 386 386 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 389 389 BY SIMILARITY.
 FT METAL 408 408 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 471 471 ZINC (CATALYTIC) (BY SIMILARITY).
 FT MOD_RES 12 12 PROTON DONOR (POTENTIAL).
 FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 316 316 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 546 546 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 754 754 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 792 792 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 945 AA; 107956 MW; A93A00CB8635F574 CRC64;
 Query Match 10.3%; Score 6; DB 1; Length 945;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 23 ISMFED 28
 DB 694 ISMFED 699
 RESULT 44
 AMPE_HUMAN STANDARD; PRT; 957 AA.
 ID AMPE_HUMAN
 AC Q07075;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glutamyl aminopeptidase (EC 3.4.11.7) (EAP) (Aminopeptidase A) (APA)
 DE (Differentiation antigen Gp160).
 GN ENPEP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 242-251; 300-316 AND 369-377.
 RC TISSUE-Kidney cortex;
 RX MEDLINE=93348214; PubMed=8346219;
 RA Natus D.M., Engelstein D., Gastl G.A., Gluck L., Vidal M.J.,
 RA Morrison M., Flinstad C.L., Bander N.H., Albino A.P.;
 RT "Molecular cloning of the human kidney differentiation antigen gp160;
 RT human aminopeptidase A.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:7069-7073(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Kidney;
 RX MEDLINE=94063909; PubMed=8244382;
 RA Li L., Wang J., Cooper M.D.;
 RT "cDNA cloning and expression of human glutamyl aminopeptidase
 RT (aminopeptidase A).";
 RL Genomics 17:657-664(1993).
 CC -1- FUNCTION: APPEARS TO HAVE A ROLE IN THE CATABOLIC PATHWAY OF THE
 CC RENIN-ANGIOTENSIN SYSTEM. PROBABLY PLAYS A ROLE IN REGULATING
 CC GROWTH AND DIFFERENTIATION OF EARLY B-LINEAGE CELLS.
 CC -1- CATALYTIC ACTIVITY: Release of a N-terminal glutamate (and to a
 CC lesser extent aspartate) from a peptide.
 CC -1- COFACTOR: REPORTEDLY ZINC-DEPENDENT, REQUIRES CALCIUM FOR FULL
 CC ACTIVITY.
 CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein.

CC -1- TISSUE SPECIFICITY: EXPRESSED BY EPITHELIAL CELLS OF THE PROXIMAL
 CC TUBULE CELLS AND THE GLOMERULUS OF THE NEPHRON. ALSO FOUND IN A
 CC VARIETY OF OTHER TISSUES.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M1.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L14721; AAA35522.1; -
 CC EMBL: L12468; AAA16876.1; -
 CC PIR: A48287; A48287.
 CC MEROPS: M01.003; -.
 CC
 CC MIM: 138297; -.
 CC
 CC InterPro: IPR001930; Aladiptase.
 CC InterPro: IPR00130; Zn_MTPetase.
 CC Pfam: PF01433; Peptidase_M1; 1.
 CC PRINTS: PR00756; ALADIPTASE.
 CC PROSITE: PS00142; ZINC_PROTEASE; 1.
 KW Hydrolase; Metalloprotease; Aminopeptidase; Zinc; glycoprotein;
 KW Antigen; Transmembrane; Phosphorylation; Signal-anchor; Calcium.
 FT DOMAIN 1 17 CYTOPLASMIC (POTENTIAL)
 FT TRANSSEM 18 40 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT
 FT DOMAIN 41 957
 FT METAL 393 393 EXTRACELLULAR (POTENTIAL).
 FT ACT_SITE 394 394 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 397 397 BY SIMILARITY.
 FT METAL 416 416 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 479 479 ZINC (CATALYTIC) (BY SIMILARITY).
 FT MOD_RES 12 12 PROTON DONOR (POTENTIAL).
 FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 554 554 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 589 589 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 597 597 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 678 678 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 763 763 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 773 773 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 801 801 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 828 828 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 218 218 A -> V (IN REF. 2).
 SQ SEQUENCE 957 AA; 109244 MW; 747E39CBB57CC574 CRC64;
 Query Match 10.3%; Score 6; DB 1; Length 957;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 23 ISMFED 28
 DB 703 ISMFED 708
 RESULT 45
 PDNG_ECOLI STANDARD; PRT; 1015 AA.
 ID PDNG_ECOLI
 AC P24183; P78261;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Formate dehydrogenase, nitrate-inducible, major subunit (EC 1.2.1.2)
 DE (formate dehydrogenase-N alpha subunit) (FDH-N alpha subunit)
 DE (anaerobic formate dehydrogenase major subunit).
 GN PDNG OR B1474.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12:
 RX MEDLINE=92042178; PubMed=1834669;
 RA Berg B.L., Li J., Heider J., Stewart V.;
 RT "Nitrate-inducible formate dehydrogenase in Escherichia coli K-12. I.
 RT Nucleotide sequence of the fdnCHI operon and evidence that opal (UGA)
 RT encodes selenocysteine.";
 RL J. Biol. Chem. 266:22380-22385(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12:
 RX MEDLINE=97251357; PubMed=9097039;
 RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
 RA Kasai H., Kashimoto K., Kimura S., Kitagawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
 RA Sempel G., Seki Y., Sivasubraman S., Tagami H., Takeda J.,
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horuchi T.;
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 28.0-40.1 min region on the linkage map.";
 RL DNA Res. 3:363-377(1996).
 CC -1- FUNCTION: FORMATE DEHYDROGENASE ALLOWS E. COLI TO USE FORMATE AS
 CC MAJOR ELECTRON DONOR DURING ANAEROBIC RESPIRATION, WHEN NITRATE IS
 CC USED AS ELECTRON ACCEPTOR. SUBUNIT ALPHA POSSIBLY FORMS THE ACTIVE
 CC SITE.
 CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
 CC -1- COFACTOR: MOLYBDENUM (MOLYBDOPTERIN) AND SELENOCYSTEINE. THE
 CC ACTIVE-SITE SELENOCYSTEINE IS ENCODED BY THE OPAL CODON, UGA.
 CC MAY BIND A 4FE-4S CLUSTER.
 CC -1- PATHWAY: ANAEROBIC NITRATE RESPIRATION.
 CC -1- SUBUNIT: FORMATE DEHYDROGENASE IS A MEMBRANE-BOUND COMPLEX, FORMED
 CC BY SUBUNITS ALPHA, BETA AND GAMMA.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- INDUCTION: BY NITRATE UNDER ANAEROBIC CONDITIONS.
 CC -1- SIMILARITY: BELONGS TO THE PROKARYOTIC MOLYBDOPTERIN-CONTAINING
 CC OXIDOREDUCTASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M75029; -; NOT ANNOTATED_CDS.
 DR EMBL: AE000244; A013438.1; -;
 DR EMBL: D90788; BAA15123.1; -;
 DR EMBL: D90789; BAA15132.1; -;
 DR PIR: JS0628; JS0628.
 DR Ecogene: Egl1227; fdng.
 DR InterPro: IPR001467; Molybdopterin.
 DR Pfam: PF00384; molybdopterin; 2.
 DR Pfam: PF01568; molybdop_binding; 1.
 DR PROSITE: PS00551; MOLYBDOPTERIN_PROK_1; 1.
 DR PROSITE: PS00490; MOLYBDOPTERIN_PROK_2; FALSE_NEG.
 DR PROSITE: PS00932; MOLYBDOPTERIN_PROK_3; 1.
 KW Oxidoreductase; Molybdenum; Selenocysteine; Selenium; NAD;
 KW Iron-sulfur; 4fe-4s; Complete proteome.
 FT METAL 50 50 IRON-SULFUR (4FE-4S) (BY SIMILARITY).

Query Match	Score 6: DB 1; Length 1015;
Best Local Similarity 100.0%; Pred. No. 71;	
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
db 703 EDLYDA 708	
27 EDLYDA 32	
53 METAL 53 IRON-SULFUR (4FE-4S) (BY SIMILARITY).	
57 METAL 57 IRON-SULFUR (4FE-4S) (BY SIMILARITY).	
92 METAL 92 IRON-SULFUR (4FE-4S) (BY SIMILARITY).	
196 SE_CYS 196	
96 CONFLICT 96	
484 CONFLICT 491	
941 CONFLICT 941	
1015 AA; 112916 MW; 0447BDD527912A7 CRC64;	
SEQUENCE	

Search completed: November 5, 2002, 05:25:27
 Job time : 31 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 5, 2002, 04:24:19 ; Search time 1796 Seconds
(without alignments)
675.801 Million cell updates/sec

Title: US-09-727-892A-99

Perfect score: 58

Sequence: 1 MRRKYTVLLYCDKIKGHFP.....YEVNLFTRKYAYIIEYIKEL 58

Scoring table:

OLIGO
Xgapop 60.0 , Ygapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1797656 segs, 10463268293 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3592925

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09727892/runat_01112002_185930_4894/app_query.fasta.1.199
-DB=genembl -QFMT=fastap -SUFFIX=NAolig.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCMATCH=200 -THR SCORE=quality -THR MIN=1 -ALIGN=45 -MODE=LOCAL -OUTFMT=ptc
-NOR=ext -HEARSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09727892.ecgn.1.1182-@runat_01112002_185930_4894 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NES_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Genembl:*
1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlggo_inv:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	13.8	36215	9	AC025728 Homo sapi
2	8	13.8	93942	9	AC006009 Homo sapi
3	8	13.8	101765	9	AL136306 Human DNA
4	8	13.8	106582	9	AL360215 Human DNA
5	8	13.8	132790	9	HS167A14
6	8	13.8	141591	9	AC079467
7	8	13.8	150762	2	AC026246
8	8	13.8	164278	9	AC093863
9	8	13.8	171690	9	AC090539
10	8	13.8	172419	9	AC098858
11	8	13.8	178776	2	AC068217
12	8	13.8	180899	9	AC022820
13	8	13.8	188804	9	AC073125
14	8	13.8	194215	2	AC024422
15	7	12.1	396	14	HIVA8615
16	7	12.1	469	3	AF174682
17	7	12.1	509	6	AX284699
18	7	12.1	573	11	G58636
19	7	12.1	673	1	BAU10498
20	7	12.1	833	8	DB9182
21	7	12.1	909	3	GI944701
22	7	12.1	948	11	CNS06J27
23	7	12.1	1050	11	CNS06J25
24	7	12.1	1501	5	XTRFTE
25	7	12.1	1611	3	AF235667
26	7	12.1	1854	8	SCYNL248C
27	7	12.1	1852	8	AF389287
28	7	12.1	1877	8	YSCA49A
29	7	12.1	1954	8	AF311953
30	7	12.1	1994	9	HSU42303
31	7	12.1	2012	8	SCYNL249C
32	7	12.1	2067	9	AF131831
33	7	12.1	2163	9	HSU85048
34	7	12.1	2214	10	AF326555
35	7	12.1	2235	10	AF326556
36	7	12.1	2333	9	AK025775
37	7	12.1	2433	10	AF202733
38	7	12.1	2497	6	A06139
39	7	12.1	2497	14	CAPYK
40	7	12.1	2554	8	SCYLR142W
41	7	12.1	2571	8	AF103948
42	7	12.1	2742	3	AY071067
43	7	12.1	2881	8	YSCPOT1
44	7	12.1	2912	10	AF208023
45	7	12.1	2912	10	MMU297397

ALIGNMENTS

RESULT 1
AC025728
LOCUS AC025728 36215 bp DNA linear PRI 07-OCT-2000
DEFINITION Homo sapiens PAC clone RP5-884M6 from 7, complete sequence.
ACCESSION AC025728
VERSION AC025728.4 GI:10047913
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
99063792
2 (bases 1 to 36215)
Stampehl,M., Maupin,R., Haakenson,B. and Atkins,V.
The sequence of Homo sapiens PAC clone RP5-884M6
Unpublished
JOURNAL
AUTHORS
TITLE
JOURNAL
4 (bases 1 to 36215)
Waterston,R.H.
Submitted (13-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 36215)
Waterston,R.H.
Submitted (10-SEP-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
5 (bases 1 to 36215)
Waterston,R.H.
Submitted (07-OCT-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Sep 10, 2000 this sequence version replaced gi:7940370.

Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@watson.wustl.edu

Summary Statistics
Center project name: H_DU0884M06

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see
http://www.nhgri.nih.gov/DIR/CTB/CHR7, send
mailto:egreen@ngri.nih.gov, or see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
This clone was derived from human PAC library RPT-5, prepared by Pieter de Jong and coworkers at the Roswell Park Cancer Institute (http://dacpac.med.buffalo.edu) using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor.
The clone may be obtained either from Genome Systems, Inc. (http://www.genomesystems.com) or Research Genetics, Inc. (http://www.resgen.com); or from Pieter de Jong.
VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP5-892G19, 200 bp overlap; the

clone sequenced to the right is CTB-133P21, 200 bp overlap. Actual start of this clone is at base position 25879 of RP5-892G19; actual end is at base position 1131 of CTB-133P21.

FEATURES
SOURCE
1..36215
location/Qualifiers

repeat_region
/rpt_family="Alu"
673..1220
repeat_region
/rpt_family="L1"
1227..1319
repeat_region
/rpt_family="Alu"
1320..1719
repeat_region
/rpt_family="ERV1"
1720..1805
repeat_region
/rpt_family="Alu"
1831..2172
repeat_region
/rpt_family="ERV"
2213..3695
repeat_region
/rpt_family="L1"
4331..4636
repeat_region
/rpt_family="Alu"
5624..6215
repeat_region
/rpt_family="L1"
7321..7520
repeat_region
/rpt_family="L1"
7530..7843
repeat_region
/rpt_family="Alu"
7844..8162
repeat_region
/rpt_family="L1"
8163..8607
repeat_region
/rpt_family="ERV1"
8609..8784
repeat_region
/rpt_family="L1"
8785..9093
repeat_region
/rpt_family="Alu"
9094..9590
repeat_region
/rpt_family="L1"
9616..10084
repeat_region
/rpt_family="ERV1"
10085..10167
repeat_region
/rpt_family="L1"
10175..10969
repeat_region
/rpt_family="L1"
10970..11363
repeat_region
/rpt_family="ERV1"
11364..11652
repeat_region
/rpt_family="L1"
11868..12207
repeat_region
/rpt_family="L2"
12216..12278
repeat_region
/rpt_family="MERL_type?"
12282..12822
repeat_region
/rpt_family="L2"
14623..14813
repeat_region
/rpt_family="MERL_type"
15547..15703
repeat_region
/rpt_family="ERV1"
15714..15930
repeat_region
/rpt_family="Alu"
16684..17389
repeat_region
/rpt_family="L1"
18506..19017
repeat_region
/rpt_family="ERV1"
19032..19372
repeat_region
/rpt_family="MALR"
19456..19762
repeat_region
/rpt_family="Alu"

```

repeat_region      19791..20241      /rpl_family="ERV"
repeat_region      20247..20474      /rpl_family="CRL"
repeat_region      21189..21888      /rpl_family="ERVK"
repeat_region      21939..22021      /rpl_family="MIR"
repeat_region      22022..22400      /rpl_family="ERV1"
repeat_region      22401..22494      /rpl_family="MIR"
repeat_region      22571..23006      /rpl_family="MALR"
repeat_region      23720..24022      /rpl_family="Alu"
repeat_region      24273..24919      /rpl_family="CRL"
repeat_region      25019..25214      /rpl_family="Alu"
repeat_region      25224..25395      /rpl_family="MIR"
repeat_region      25852..25898      /rpl_family="ERV1"
repeat_region      25901..25946      /rpl_family="CRL"
repeat_region      26246..26608      /rpl_family="L1"
repeat_region      26634..26739      /rpl_family="L1"
repeat_region      27449..27587      /rpl_family="CRL"
repeat_region      28346..28580      /rpl_family="MIR"
repeat_region      28663..29306      /rpl_family="MER2_type"
repeat_region      30684..30973      /rpl_family="ERV1"
repeat_region      30974..31337      /rpl_family="L2"
repeat_region      31354..31399      /rpl_family="MALR"
repeat_region      31426..31836      /rpl_family="L2"
repeat_region      31837..32487      /rpl_family="MALR"
repeat_region      33004..33177      /rpl_family="L2"

```

```

Alignment Scores:
Pred. No.:      55.6      Length:      36215
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches:  0
Query Match:     13.79%      Indels:      0
DB:              9      Gaps:          0

```

US-09-727-892A-99 (1-58) x AC025728 (1-36215)

Oy 42 AsnleuphermLysLysTYRAla 49
 Db 3201 AATCTATTACAAAGAAATATGCT 3224

RESULT 2

```

LOCUS      AC006009      93942 bp      DNA      linear      PRI 21-DEC-1999
DEFINITION Homo sapiens PAC clone RP5-85011 from 7q31.2-q32, complete.
ACCESSION  AC006009
VERSION     AC006009.2  GI:4753278
KEYWORDS    HMG.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE
AUTHORS      1 (bases 1 to 93942)
TITLE        Sulston, J.E. and Waterston, R.
JOURNAL      Toward a complete human genome sequence
MEDLINE      Genome Res. 8 (11), 1097-1108 (1998)
99063792
REFERENCE
AUTHORS      2 (bases 1 to 93942)
TITLE        Maupin, R., Elliott, G., Bauer, C. and Lehnert, L.
JOURNAL      The sequence of Homo sapiens PAC clone RP5-85011
unpublished
REFERENCE
AUTHORS      3 (bases 1 to 93942)
TITLE        Waterston, R.H.
JOURNAL      Direct Submission
unpublished (22-NOV-1998) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE
AUTHORS      4 (bases 1 to 93942)
TITLE        Waterston, R.H.
JOURNAL      Direct Submission
Submitted (05-MAY-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE
AUTHORS      5 (bases 1 to 93942)
TITLE        Waterston, R.
JOURNAL      Direct Submission
Submitted (17-JUN-1999) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
6 (bases 1 to 93942)
REFERENCE
AUTHORS      Direct Submission
JOURNAL      Submitted (21-DEC-1999) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On May 5, 1999 this sequence version replaced gi:4204344.
COMMENT
-----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.wustl.edu
-----
Summary Statistics
Center project name: HJD0850101
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nih.gov/DIR/OTB/CHR/>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

This clone was derived from human PAC library RPI-5, prepared by Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://pacpac.med.buffalo.edu>) using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor. The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc.

(http://www.resgen.com); or from Pieter de Jong.
 VECTOR: PCYPAC2
 NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is RP5-1101C3, 200 bp overlap.
 Actual start of this clone is at base position 122171 of
 RP5-1101C3; actual end is at 93942 of RP5-85011.
 Location/Qualifiers

FEATURES
 source
 1..93942
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7q31.2-q32"
 /clone="RP5-85011"
 /clone_1id="RPC1-5"
 89..306
 repeat_region
 /rpt_family="Alu"
 380..632
 /rpt_family="L1"
 1346..1620
 /rpt_family="Retroviral"
 repeat_region
 1940..1981
 /rpt_family="AT_rich"
 2935..3362
 repeat_region
 /rpt_family="L1"
 3366..3541
 /rpt_family="L1"
 3542..3573
 repeat_region
 /rpt_family="(TAAA)n"
 3574..3855
 /rpt_family="Alu"
 3856..4738
 /rpt_family="L1"
 5050..5362
 repeat_region
 /rpt_family="L1"
 5363..5641
 repeat_region
 /rpt_family="Alu"
 5642..6174
 repeat_region
 /rpt_family="L1"
 6811..6904
 repeat_region
 /rpt_family="MER86"
 8217..8239
 /rpt_family="AT_rich"
 8367..8666
 repeat_region
 /rpt_family="Alu"
 8962..9298
 repeat_region
 /rpt_family="Alu"
 10260..10424
 repeat_region
 /rpt_family="MER1_type"
 10425..10727
 repeat_region
 /rpt_family="Alu"
 10728..10869
 repeat_region
 /rpt_family="MER1_type"
 11128..11380
 repeat_region
 /rpt_family="L2"
 11529..11676
 repeat_region
 /rpt_family="MIR"
 11964..12255
 repeat_region
 /rpt_family="Alu"
 14116..14137
 repeat_region
 /rpt_family="AT_rich"
 14329..14670
 repeat_region
 /rpt_family="L2"
 14923..15319
 repeat_region
 /rpt_family="L2"
 15389..15443
 repeat_region
 /rpt_family="L2"
 15464..15533
 repeat_region
 /rpt_family="(CATA)n"
 15847..16022
 repeat_region
 /rpt_family="Alu"
 17093..17297
 repeat_region
 /rpt_family="MIR"
 complement(18038..18356)

repeat_region
 /db_xref="GI:1113435"
 20102..20159
 /rpt_family="AT_rich"
 repeat_region
 20168..20590
 /rpt_family="MER1_type"
 repeat_region
 22057..22327
 /rpt_family="L1"
 repeat_region
 23137..23216
 /rpt_family="L1"
 repeat_region
 23217..23263
 /rpt_family="AT_rich"
 repeat_region
 23289..23330
 /rpt_family="AT_rich"
 repeat_region
 23647..23931
 /rpt_family="Alu"
 repeat_region
 24011..24041
 /rpt_family="AT_rich"
 repeat_region
 24303..24928
 /rpt_family="L1"
 repeat_region
 25690..25711
 /rpt_family="AT_rich"
 repeat_region
 26770..26789
 /rpt_family="(CAA)n"
 repeat_region
 27146..27172
 /rpt_family="AT_rich"
 repeat_region
 29508..29808
 /rpt_family="Alu"
 repeat_region
 29828..30626
 /rpt_family="L1"
 repeat_region
 30952..31068
 /rpt_family="L1"
 repeat_region
 32261..32546
 /rpt_family="Alu"
 repeat_region
 32995..33119
 /rpt_family="MER2_type"
 repeat_region
 33199..33388
 /rpt_family="MER2_type"
 repeat_region
 33574..33603
 /rpt_family="(TAAA)n"
 repeat_region
 33604..33884
 /rpt_family="Alu"
 repeat_region
 34210..34360
 /rpt_family="L2"
 repeat_region
 34977..35012
 /rpt_family="(GA)n"

Alignment Scores:

pred. No.: 119
 Score: 8.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 13.79%
 DB: 9
 Gaps: 0

US-09-727-892a-99 (1-58) x AC006009 (1-93942)

QY 4 LYSTYLSTHrValleuLeuTyr 11
 |||||
 Db 16974 AAGTACAAACAGTCTACTAT 16951

RESULT 3

LOCUS

DEFINITION

AL136306 101765 bp. DNA linear prt 21-DEC-2000
 Human DNA sequence from clone RP3-334F4 on chromosome 6 Contains
 ESTs, STSs and GSSs. Contains a LAMR1 (laminin receptor 1,
 ribosomal protein SA) pseudogene and an RPL10 (ribosomal protein
 L10) pseudogene, complete sequence.

ACCESSION

AL136306

VERSION

AL136306.11 GI:10045289

KEYWORDS

HTG; laminin receptor; LAMR1; ribosomal protein; RPL10.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 101765)
AUTHORS Johnson,C.
TITLE Direct Submission
JOURNAL Submitted (21-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

COMMENT

On Sep 9, 2000 this sequence version replaced gi:5588478. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission only a small overlap as described above. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/Ref/Chr6>

IMPORTANT: This sequence is not the entire insert of clone RP3-334F4. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP3-334F4 is at 1 in this sequence. The true right end of clone RP3-510B21 is at 101666 in this sequence. The true right end of clone RP11-47D15 is at 22579 in this sequence. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. RP3-334F4 is from the library RPc1-3 constructed by the group of Pletier de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pCYPAC2.

FEATURES

source

location/Qualifiers
1..101765
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP3-334F4"
/clone.lib="RPC1-3"
13..290
repeat_region
/note="AluJ6 repeat: matches 6. .284 of consensus"
297..352
repeat_region
/note="28 copies 2 mer ag 80% conserved"
358..560
repeat_region
/note="MIR repeat: matches 51. .238 of consensus"
1635..1707
repeat_region
/note="L2 repeat: matches 2606. .2675 of consensus"
2470..2667
repeat_region
/note="L1ME3 repeat: matches 5666. .5896 of consensus"
3417..3963
misc_feature
/note="match: GSS: Em:AQ208268"
3606..4123
misc_feature
/note="match: GSS: Em:A0513080"
3969..4235
repeat_region
/note="L1P3 repeat: matches 5156. .5431 of consensus"
5300..5685
repeat_region
/note="MSTA repeat: matches 2. .426 of consensus"
5826..6235
repeat_region
/note="M172F repeat: matches 3. .407 of consensus"
6248..6909
misc_feature
/note="match: STS: Em:G58052
match: GSS: Em:AQ348294"
7040..7427
repeat_region
/note="M17E repeat: matches 180. .568 of consensus"
complement(8704..9014)
misc_feature
/note="match: STS: Em:G02980"
9461..9617
repeat_region

misc_feature
/note="MIR repeat: matches 28. .192 of consensus"
complement(10805..11198)
/note="match: GSS: Em:AQ181524"
10840..11244
repeat_region
/note="MSTA repeat: matches 2. .426 of consensus"
11203..11718
misc_feature
/note="match: GSS: Em:A0670295"
11420..11472
repeat_region
/note="M17D3 repeat: matches 497. .553 of consensus"
11488..11515
repeat_region
/note="14 copies 2 mer gt 94% conserved"
11544..11886
repeat_region
/note="M17D2 repeat: matches 17. .364 of consensus"
11958..12066
repeat_region
/note="AluJ/FLAM repeat: matches 5. .83 of consensus"
12067..12097
repeat_region
/note="BC200 repeat: matches 5. .131 of consensus"
12198..12485
repeat_region
/note="AluX repeat: matches 3. .291 of consensus"
12580..13685
repeat_region
/note="L1ME1 repeat: matches 5027. .617 of consensus"
13686..16339
repeat_region
/note="T1GGER2 repeat: matches 1. .2708 of consensus"
16539..16862
repeat_region
/note="L1ME2 repeat: matches 5759. .5972 of consensus"
complement(16925..17808)
/gene="dJ334F4.1"
complement(16925..17808)
/gene="dJ334F4.1"
/note="dJ334F4.1 (L1MRL1 (laminin receptor 1, ribosomal protein S4) pseudogene)
match: CDNAS: Em:AF116719 Em:D52524 Em:D28372 Em:M64923
Em:J03799 Em:X06406 Em:X61156 Em:Z22749 Em:X15005
Em:S37431 Em:M41419 Em:L16589 Em:U25771 Em:AF140348
Em:M27798 Em:U02870
match: ESTs: Em:AW628356 Em:AW518468 Em:AA088403
Em:AW589551 Em:AA622366 Em:AW38869 Em:AA155961 Em:W52358
Em:W19382 Em:AA604468 Em:AT446603 Em:AA835456 Em:AA027598
Em:AA932739 Em:N89876 Em:AA826786 Em:AA559967 Em:A1528478
Em:AW245912 Em:AW250548 Em:A1859973 Em:R62252 Em:AA328537
Em:AW249238 Em:AW389260 Em:AW398261 Em:AA204832
Em:AW661917 Em:AA530840 Em:AA082354 Em:AA389264
Em:AA522580 Em:AW627657 Em:AA730904 Em:AW619792
Em:AA085660 Em:TA48292 Em:AA649800 Em:A1921017 Em:AA691364
Em:AA709355 Em:AA592149 Em:A1554001 Em:AA411023
Em:AA653786 Em:AA815788 Em:A50707 Em:AW301438 Em:AA250715
Em:AW411348 Em:A1310230 Em:AA224522 Em:W10275 Em:AA205815
Em:T65950 Em:C06500 Em:AA303589 Em:AA606984 Em:AA858149
Em:AA377669 Em:AW117640 Em:A66334 Em:AW406804 Em:AA213123
Em:AW249171 Em:A1469759 Em:AA582040 Em:AA634053
Em:AA133653 Em:AA505038 Em:AA928220 Em:AW732895
Em:AW575980 Em:A1744200 Em:AA352263 Em:AA630178
Em:AA655423 Em:AW327989 Em:AW411044 Em:F27335 Em:R44109
Em:A1569634 Em:TT1050 Em:AW68052 Em:AW606392 Em:AA249345
Em:AA205272 Em:AA452228 Em:T47352 Em:R29157 Em:AA332674
Em:T94804 Em:AW769097 Em:AA488405 Em:A1031939 Em:A1625390
Em:AA792475 Em:AA976493 Em:AW406025 Em:AA205440
Em:AA575603 Em:AA312528 Em:AA406379 Em:A49586 Em:A1830014
Em:AW276487 Em:AA096074 Em:T65905 Em:AA555093 Em:AA375194
Em:AA1199334 Em:A165657 Em:A1082474 Em:A1088216
Em:AA228994 Em:AW327455 Em:AW247663 Em:AA654080
Em:AA109862 Em:AA619335 Em:AA486164 Em:A1628005
Em:A1719600 Em:AA903965 Em:AA085956 Em:W19848 Em:AA65658
Em:AA08757 Em:A1929620 Em:AA089518 Em:A1663158
Em:AW410265 Em:AW246303 Em:AW515363 Em:AA629810
Em:AA246860 Em:AW404829 Em:N93605 Em:AA714360 Em:AA670450
Em:AA066667 Em:AA229109 Em:AA131199 Em:AA2646788
Em:AA129319 Em:AA497505 Em:AA689621 Em:T56857 Em:AA140111
Em:AA798646 Em:AA489489 Em:AW615423 Em:N57410 Em:AW576041
match: proteins: Sw:P46769 Sw:P14206 Tr:O29231 Tr:O80377
Sw:P50890 Tr:Q9Y7L8 Sw:P38979 Sw:P08865 Sw:P38980
Tr:O9W583 Sw:P38981 Sw:P38982 Sw:P38983 Sw:P38984
Sw:P26452 Sw:P46770 Sw:P46771 Tr:Q9U5P8

```

/codon_start=1
/pseudo
/evidence=not_experimental
complement(117020..17141)
/gene="dJ334f4.1"
/notes="match: STS: Em:G05566 Em:G59752"
complement(117326..17640)
/gene="dJ334f4.1"
/notes="match: GSS: Em:B31090"
17888..18169
/notes="LIME2 repeat: matches 5484..5766 of consensus"
18714..19632
/notes="MER21A repeat: matches 22..901 of consensus"
19280..19737
/notes="match: GSS: Em:AQ881491"
19306..19819
/notes="match: GSS: Em:AQ332579"
19686..19841
/notes="MIR repeat: matches 50..210 of consensus"
20995..21174
/notes="MIR repeat: matches 69..256 of consensus"
21260..21528
/notes="LTR16A repeat: matches 18..289 of consensus"
21531..21635
/notes="MIR repeat: matches 65..172 of consensus"
complement(22146..22579)
/notes="match: STS: Em:C51546
match: GSS: Em:AQ200401"
24860..25612
/notes="LIME2 repeat: matches 5319..6092 of consensus"
25772..27223
/notes="LIP3 repeat: matches 569..1870 of consensus"
complement(26901..27403)
/notes="match: GSS: Em:AQ040587"
complement(27004..27397)
/notes="match: GSS: Em:AQ113410"
complement(27060..27597)
/notes="match: GSS: Em:AQ533137"
27170..27256
/notes="match: GSS: Em:AQ380004"
27261..27745
/notes="match: GSS: Em:AQ380004"
27582..27911
/notes="match: GSS: Em:AQ380004"
repeat_region

Alignment Scores:
Pred. No.: 127 Length: 101765
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.79% Indels: 0
Gaps: 0

US-09-727-892a-99 (1-58) x AL136306 (1-101765)

QY 4 LysTyrIystrValIleuLeuTyr 11
Db 20475 AAATATAAACTGCTGCTGTA 20498

RESULT 4
AL360215/c 106582 bp DNA linear PRI 09-MAY-2001
LOCUS AL360215 Human DNA sequence from clone RP11-535D18 on chromosome 6 Contains
DEFINITION STSs, GSSs and a CpG island, complete sequence.
ACCESSION AL360215
VERSION AL360215.15 GI:109444203
KEYWORDS HMG; CpG island.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 106582)
AUTHORS Chapman,J.
TITLE Direct Submision

```

JOURNAL Submitted (06-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequests@sanger.ac.uk
 On Oct 20, 2000 this sequence version replaced gi:10803289.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6>

IMPORTANT: This sequence is not the entire insert of clone RP11-535D18 it may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true right end of clone RP11-535D18 is at 106582 in this sequence. The true right end of clone RP11-535D18 is at 100 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats: all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP11-535D18 is from the library RPICT-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

FEATURES

SOURCE

```

location/Qualifiers
1..106582
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-535D18"
/clone_1bp="RPICT-11.2"
1..4421
/notes="LIP47 repeat: matches 1687..6140 of consensus"
4486..4992
/notes="LTR1A2-internal repeat: matches 1..512 of consensus"
4993..5367
/notes="LTR1A2 repeat: matches 1..374 of consensus"
5368..6133
/notes="L2 repeat: matches 1553..2378 of consensus"
6159..6447
/notes="LIME3A repeat: matches 5680..5959 of consensus"
10249..10312
/notes="MIR repeat: matches 87..150 of consensus"
complement(10776..11227)
/notes="match: STS: Em:G08565"
complement(10836..10998)
/notes="match: STS: Em:G07938"
complement(10840..11069)
/notes="match: STS: Em:G07938"
10883..11079
/notes="match: STS: Em:A2412555"
complement(10899..11075)
/notes="match: GSS: Em:AQ96755"
10904..11069
/notes="match: GSS: Em:AQ35323"
complement(10904..11075)
/notes="match: GSS: Em:AQ377847 Em:A253034"
complement(10904..11046)
/notes="match: STS: Em:LI6408"
complement(10904..11034)
/notes="match: GSS: Em:AL230661"

```

misc_feature	10904..11015	/note="match: GSS: Em:AQ0540280"
misc_feature	complement(10904..11018)	
misc_feature	/note="match: SFS: Em:L30694"	
misc_feature	10904..11031	/note="match: SFS: Em:G09117"
misc_feature	complement(10905..11058)	
misc_feature	/note="match: GSS: Em:A2430872"	
misc_feature	complement(10906..11079)	
misc_feature	/note="match: GSS: Em:A2040457"	
misc_feature	10906..11033	
misc_feature	/note="match: GSS: Em:A2783933"	
misc_feature	10906..11075	
misc_feature	/note="match: GSS: Em:AQ056687"	
misc_feature	10906..11084	/note="match: GSS: Em:AQ057863"
misc_feature	complement(10906..11079)	
misc_feature	/note="match: SFS: Em:G10534"	
misc_feature	10906..11069	
misc_feature	/note="match: GSS: Em:AL231617"	
misc_feature	10907..11055	
misc_feature	/note="match: GSS: Em:A2370192"	
misc_feature	complement(10913..11035)	
misc_feature	/note="match: GSS: Em:A251476"	
misc_feature	10913..11034	
misc_feature	/note="match: GSS: Em:A2065198"	
misc_feature	complement(10913..11008)	
misc_feature	/note="match: SFS: Em:G07837 Em:G34878"	
misc_feature	complement(10913..11013)	
misc_feature	/note="match: SFS: Em:117772"	
misc_feature	complement(10914..11075)	
misc_feature	/note="match: GSS: Em:A2122363"	
misc_feature	complement(10916..11023)	
misc_feature	/note="match: GSS: Em:B33399"	
misc_feature	complement(10919..11073)	
misc_feature	/note="match: SFS: Em:G27278"	
repeat_region	10919..10934	
repeat_region	/note="9 copies 4 mer tail 100% conserved"	
misc_feature	complement(10924..11050)	
misc_feature	/note="match: SFS: Em:L18725"	
misc_feature	complement(10925..11060)	
misc_feature	/note="match: SFS: Em:G09149"	
misc_feature	10928..11059	
misc_feature	/note="match: SFS: Em:G09635"	
misc_feature	10950..11069	
misc_feature	/note="match: SFS: Em:G09622"	
misc_feature	complement(10961..11075)	
misc_feature	/note="match: GSS: Em:A2271663"	
misc_feature	10962..11078	
misc_feature	/note="match: SFS: Em:G07836"	
misc_feature	10972..11084	
misc_feature	/note="match: SFS: Em:G10403"	
misc_feature	10983..11075	
misc_feature	/note="match: SFS: Em:G09625"	
repeat_region	12025..12057	
repeat_region	/note="ME5A repeat: matches 21..49 of consensus"	
repeat_region	12058..12222	
repeat_region	/note="ME45 repeat: matches 1..176 of consensus"	
repeat_region	12223..12373	
repeat_region	/note="ME5A repeat: matches 49..187 of consensus"	
repeat_region	12458..12829	
repeat_region	/note="match: GSS: Em:AQ32075"	
repeat_region	13382..13740	
repeat_region	/note="HEH repeat: matches 1..369 of consensus"	
repeat_region	14457..14925	
repeat_region	/note="LMB2 repeat: matches 5666..6143 of consensus"	
repeat_region	15702..15959	
repeat_region	/note="AluSq repeat: matches 37..292 of consensus"	
repeat_region	17847..18161	
repeat_region	/note="AluSq repeat: matches 1..312 of consensus"	
repeat_region	18162..18296	
repeat_region	/note="AluSq repeat: matches 1..135 of consensus"	
repeat_region	18298..19116	

```

repeat_region /note="L2 repeat: matches 1568. .2487 of consensus"
20829. .2114
/note="L1M7 repeat: matches 5961. .6284 of consensus"
repeat_region 21181. .2207
/note="PRT5 repeat: matches 743. .1561 of consensus"
misc_feature 21194. .21748
/note="Cpg island"
/evidence-not_experimental
repeat_region 22016. .2276
/note="LTR12 repeat: matches 1. .671 of consensus"
repeat_region 23074. .23156
/note="MIR repeat: matches 171. .252 of consensus"
repeat_region 23113. .23164
/note="L2 repeat: matches 2651. .2702 of consensus"
repeat_region 23551. .23834
/note="Aluab repeat: matches 2. .306 of consensus"
repeat_region 23860. .23943
/note="42 copies 2 mer tt 67% conserved"
repeat_region 24341. .24430
/note="MIR repeat: matches 9. .98 of consensus"
repeat_region 24585. .24967
/note="W81A repeat: matches 1. .426 of consensus"
repeat_region 25091. .25576
/note="L2 repeat: matches 379. .895 of consensus"
repeat_region 25582. .26028
/note="LFR32 repeat: matches 15. .465 of consensus"
repeat_region 26046. .26742
/note="L2 repeat: matches 903. .2744 of consensus"
repeat_region 27796. .27875
/note="MEB94 repeat: matches 54. .133 of consensus"
repeat_region 28099. .28393
/note="AluSq repeat: matches 3. .297 of consensus"
repeat_region 28670. .28709
/note="20 copies 2 mer tg 92% conserved"
repeat_region 29025. .29172
/note="L1A6 repeat: matches 6125. .6284 of consensus"
repeat_region 29199. .29306
/note="MEB81 repeat: matches 2. .111 of consensus"
misc_feature complement(29542. .29913)
/note="match: GSS: Em: A0009707"
repeat_region 30351. .30466

```

```

Alignment Scores:
Pred. No.: 132
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 13.79%
DB: 9

US-09-727-892A-99 (1-58) x AL360215 (1-106582)
OY 3 ArgLysTyrLysThrValLeuLeu 10
|||||
Db 25471 AGAAGTATAAACAGTCTCTTG 25448

RESULT 5
HS167A14/c
LOCUS HS167A14 132790 bp DNA linear PRI 23-NOV-1999
DEFINITION Human DNA sequence from PAC 167A14 on chromosome 6q27. Contains
ESTs, CPG island and STRs.
ACCESSION Z94721
VERSION 294721.1 GI:2462374
KEYWORDS 6q27; CPG island.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 132790)
Phillips, S.
Direct Submission
Submitted (25-SEP-1997) Chromosome 6 Project Group
(http://www.sanger.ac.uk/HGP/Chr6/) Sanger Centre, Hinxton,

```

COMMENT

Cambridgeshire, CB10 1SA, UK. E-mail enquires:
humquery@sanger.ac.uk Clone request: clonerequest@sanger.ac.uk
On Oct 2, 1997 this sequence version replaced gi:2052030.
IMPORTANT: This sequence is the entire insert of clone 167A14.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together
with a note of the overlapping clone name. Note that the variations
annotated may not be found in the sequence submission corresponding
to the overlapping clone as we submit sequences with only a small
overlap
as described above.
This sequence was generated from part of bacterial clone contigs of
human chromosome 6, constructed by the Sanger Centre chromosome 6
mapping group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr6/>
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key
The true left end of clone 167A14 is at 1 in this sequence. The
true right end of this clone is at 132790.
167A14 is from the library RPC11 constructed at the Roswell Park
Cancer
Institute by the group of Pieter de Jong.
For further details see <http://Daacpac.med.buffalo.edu/>.

FEATURES

source

Location/Qualifiers
1. .132790
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6q27"
/clone="RP1-167A14"
/clone_lib="RPC1-1"
1183. .1450
/note="MER42c repeat: matches 1158. .879 of consensus"
2042. .2089
/note="2 copies of 24 mer 96 & conserved"
3211. .3512
/note="AluSq repeat: matches 1. .302 of consensus"
3903. .3976
/note="MER33 repeat: matches 320. .246 of consensus"
3981. .4266
/note="AluSp repeat: matches 1. .303 of consensus"
4289. .4510
/note="MER33 repeat: matches 231. .1 of consensus"
7421. .7703
/note="LIME2 repeat: matches 621. .902 of consensus"
8106. .8514
/note="MLT1C repeat: matches 445. .3 of consensus"
10066. .10190
/note="MIR repeat: matches 15. .136 of consensus"
11611. .11705
/note="MER5A repeat: matches 100. .187 of consensus"
12002. .12295
/note="AluIb repeat: matches 298. .1 of consensus"
13014. .13314
/note="AluSg repeat: matches 300. .1 of consensus"
13577. .13880
/note="AluSg repeat: matches 303. .1 of consensus"
14093. .19313
/note="LI repeat: matches 177. .5390 of consensus"
19164. .20058
/note="LIP2 repeat: matches 1. .891 of consensus"
22511. .22823
/note="MLT1A repeat: matches 5. .310 of consensus"
22971. .23170
/note="MIR repeat: matches 262. .46 of consensus"
23424. .23513
/note="MIR repeat: matches 205. .109 of consensus"

repeat_region 25020. .25317
/note="AluX repeat: matches 1. .300 of consensus"
repeat_region 25858. .26807
/note="LI repeat: matches 4064. .3061 of consensus"
repeat_region 27012. .27974
/note="MER25 repeat: matches 1166. .2136 of consensus"
repeat_region 27438. .30305
/note="LI repeat: matches 1103. .3927 of consensus"
repeat_region 30319. .30513
/note="AluSc repeat: matches 97. .292 of consensus;
Incomplete repeat"
repeat_region 30518. .30556
/note="AluY repeat: matches 1. .135 of consensus;
Incomplete repeat"
repeat_region 30691. .32187
/note="LI repeat: matches 3940. .5390 of consensus"
repeat_region 32030. .32925
/note="LIP3 repeat: matches 1. .897 of consensus"
repeat_region 33617. .33774
/note="AluX repeat: matches 301. .144 of consensus;
Incomplete repeat"
repeat_region 33776. .33898
/note="AluSq repeat: matches 7. .129 of consensus;
Incomplete repeat"
repeat_region 33902. .34109
/note="LIMB8 repeat: matches 856. .642 of consensus"
repeat_region 34107. .34280
/note="LIMB7 repeat: matches 609. .416 of consensus"
repeat_region 34344. .35096
/note="MER21B repeat: matches 1. .790 of consensus"
repeat_region 35115. .35514
/note="LIMB6 repeat: matches 417. .3 of consensus"
repeat_region 35714. .35872
/note="MLT2A repeat: matches 1. .160 of consensus"
repeat_region 35875. .37058
/note="MLT2-internal repeat: matches 4224. .5402 of
consensus"
repeat_region 37062. .37363
/note="AluY repeat: matches 1. .301 of consensus"
repeat_region 37364. .37610
/note="MLT2-internal repeat: matches 5423. .5669 of
consensus"
repeat_region 37611. .38077
/note="MLT2A repeat: matches 1. .451 of consensus"
repeat_region 38080. .40446
/note="LI repeat: matches 5035. .2631 of consensus"
repeat_region 40858. .40938
/note="MER21B repeat: matches 790. .718 of consensus"
repeat_region 40982. .41670
/note="MER21B repeat: matches 713. .10 of consensus"
repeat_region 42480. .42680
/note="MER4B repeat: matches 1. .191 of consensus"
repeat_region 42684. .42974
/note="AluSq repeat: matches 294. .3 of consensus"
repeat_region 43293. .43682
/note="MER4B repeat: matches 220. .574 of consensus"
repeat_region 44819. .45109
/note="AluX repeat: matches 1. .291 of consensus"
repeat_region 45295. .45557
/note="MER2 repeat: matches 345. .49 of consensus"
repeat_region 46222. .46521
/note="AluX repeat: matches 1. .301 of consensus"
repeat_region 47819. .48112
/note="AluX repeat: matches 1. .302 of consensus"
repeat_region 48282. .48521
/note="match STS G19334"
misc_feature 48282. .48751
/note="match: 3' EST T78087"
repeat_region 49468. .50157
/note="LIR8 repeat: matches 1. .676 of consensus"
repeat_region 50206. .50505
/note="AluX repeat: matches 302. .1 of consensus"
repeat_region 50603. .50976

```
/note="THEIC repeat: matches 371. .1 of consensus"
repeat_region 51051. .51789
/note="LIMB7 repeat: matches 164. .914 of consensus"
repeat_region 52449. .52741
/note="ALUdB repeat: matches 293. .13 of consensus"
repeat_region 52744. .52888
/note="ALUdJ repeat: matches 145. .4 of consensus;
incomplete repeat"
repeat_region 53445. .53643
/note="MIR repeat: matches 262. .51 of consensus"
repeat_region 55249. .55550
/note="AluY repeat: matches 298. .1 of consensus"
repeat_region 55578. .55897
/note="AluX repeat: matches 1. .302 of consensus"
repeat_region 56788. .57098
/note="MIR1A1 repeat: matches 365. .7 of consensus"
repeat_region 57318. .57622
/note="AluSg repeat: matches 1. .303 of consensus"
repeat_region 58159. .58184
/note="13 copies of 2 mer 96 % conserved"
repeat_region 58884. .59158
/note="AluSg repeat: matches 1. .280 of consensus;
incomplete repeat"
prim_transcript 60683. .61099
/note="match: multiple ESTs; match: AA292127 N39149
AA268645"
prim_transcript complement(60783. .61099)
/note="match: 3' EST AA292126; paired with EST AA292127
matching this clone"
repeat_region 61163. .61571
/note="L1PB1 repeat: matches 170. .576 of consensus"
repeat_region 61572. .61872
/note="AluSx repeat: matches 1. .302 of consensus"
repeat_region 61873. .62190
/note="L1PB1 repeat: matches 575. .901 of consensus"
repeat_region 63896. .64194
/note="AluSx repeat: matches 2. .302 of consensus"
repeat_region 64995. .65081
/note="MIR repeat: matches 262. .170 of consensus"

Alignment Scores:
Pred. No.: 158 Length: 132790
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.79% Indels: 0
DB: 9 Gaps: 0

US-09-727-892a-99 (1-58) x HS167A14 (1-132790)
QY 3 ArglystYrlystHrVallieu 10
Db 41807 AGGAATATAGACTGTCTCTTG 41784

RESULT 6
AC079467 141591 bp DNA linear PRI 30-MAR-2001
LOCUS Homo sapiens chromosome 5 clone RP11-357D18, complete sequence.
ACCESSION AC079467.2 GI:13489158
VERSION AC079467.2
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 141591)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 141591)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-2000) Production Sequencing Facility, DOE Joint
```

```
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 141591)
DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE 1 (bases 1 to 141591)
AUTHORS Direct Submission
TITLE Submitted (30-MAR-2001) DOE Joint Genome Institute, 2800 Mitchell
JOURNAL Drive, Walnut Creek, CA 94598, USA
COMMENT On Mar 30, 2001 this sequence version replaced gi:9964832.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.2.
FEATURES
SOURCE
1. 141591
/organism="Homo sapiens"
/db_xref="taxon:9606"
/contig="RP11-357D18"
/clone="RP11-357D18"
BASE COUNT 45445 a 26359 c 25389 g 44398 t
ORIGIN

Alignment Scores:
Pred. No.: 166 Length: 141591
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.79% Indels: 0
DB: 9 Gaps: 0

US-09-727-892a-99 (1-58) x AC079467 (1-141591)
QY 2 GUAArglystYrlystHrVallieu 9
Db 2315 GAAAGGAATATCAAACTGTATTA 2338

RESULT 7
AC026246/c 150762 bp DNA linear HTG 06-APR-2000
LOCUS Homo sapiens clone RP11-62N21, WORKING DRAFT SEQUENCE, 13 unordered
pieces.
ACCESSION AC026246.2 GI:7465153
VERSION AC026246.2
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 150762)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome, clone RP11-62N21
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 150762)
AUTHORS Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, D.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-pierre, N., Grant, G., Hagos, B., Heathford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehotzky, J.,
Levine, R., Liu, C., Liu, G., Locke, R., MacDonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGuck, A., McKernan, K., McPheters, R.,
Melchior, J., Menous, L., Milnova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Olivier, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rotman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., J.,
```

TITLE
JOURNAL

Young, G., Zainoun, J., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (21-MAR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 6, 2000 this sequence version replaced gi:7272146.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
 Genome Center

Center code: WtBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information

Center Project name: L5157
 Center clone name: 62.N.21

Summary Statistics

Sequencing vector: M13; M77815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 14414 bases at least Q40
 Consensus quality: 14748 bases at least Q30
 Consensus quality: 14867 bases at least Q20
 Insert size: 151000; agarose-1p
 Insert size: 149562; sum-of-coverage
 Quality coverage: 4.6 in Q20 bases; agarose-1p
 Quality coverage: 4.6 in Q20 bases; sum-of-coverage

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 13 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 2064: contig of 2064 bp in length
 * 2065 2164: gap of 100 bp
 * 2165 5722: contig of 3558 bp in length
 * 5723 5822: gap of 100 bp
 * 5823 10815: contig of 4993 bp in length
 * 10816 10915: gap of 100 bp
 * 10916 16014: contig of 5099 bp in length
 * 16015 16114: gap of 100 bp
 * 16115 22697: contig of 6583 bp in length
 * 22698 22797: gap of 100 bp
 * 22798 29224: contig of 6427 bp in length
 * 29225 29324: gap of 100 bp
 * 29325 36736: contig of 7412 bp in length
 * 36737 36836: gap of 100 bp
 * 36837 46147: contig of 9311 bp in length
 * 46148 46247: gap of 100 bp
 * 46248 58403: contig of 12156 bp in length
 * 58404 58503: gap of 100 bp
 * 58504 70029: contig of 11526 bp in length
 * 70030 70129: gap of 100 bp
 * 70130 93712: contig of 23583 bp in length
 * 93713 93812: gap of 100 bp
 * 93813 121324: contig of 27512 bp in length
 * 121325 121424: gap of 100 bp
 * 121425 150762: contig of 29338 bp in length.
 Location/Qualifiers
 1. 150762

FEATURES
source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="RP11-62N21"
 /clone_lib="RP11-11 Human Male BAC"
 1. 2064
 /note="assembly-fragment"
 2165. 5722
 /note="assembly-fragment"
 5823. 10815
 /note="assembly-fragment"
 misc_feature
 misc_feature
 misc_feature
 misc_feature

misc_feature

10916. 16014
 /note="assembly-fragment"
 clone_end:SP6
 vector_side:left"

misc_feature

16115. 22697
 /note="assembly-fragment"
 22798. 29224
 /note="assembly-fragment"

misc_feature

29325. 36736
 /note="assembly-fragment"
 36837. 46147
 /note="assembly-fragment"

misc_feature

46248. 58403
 /note="assembly-fragment"
 58504. 70029
 /note="assembly-fragment"

misc_feature

70130. 93712
 /note="assembly-fragment"
 93813. 121324
 /note="assembly-fragment"

misc_feature

clone_end:T7
 vector_side:right"
 121425. 150762
 /note="assembly-fragment"

misc_feature

25212 c 25481 g 47457 t 1201 others
 BASE COUNT 51411 a 25212 c 25481 g 47457 t 1201 others
 ORIGIN

misc_feature

Alignment Scores:
 Align. No.: 175
 Score: 8.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 13.79%
 DB: 2
 Gaps: 0

misc_feature

US-09-727-892a-99 (1-58) x AC026246 (1-150762)
 QY 7 ThrValLeuLeuTyrCysAspGlu 14
 Db 96059 ACAGTCCCTCATATGTGATGAA 96036

misc_feature

RESULT 8 164278 bp DNA linear PRI 13-FEB-2002
 AC093863
 LOCUS Homo sapiens chromosome 4 clone RP11-552110, complete sequence.
 DEFINITION AC093863 AC023631
 ACCESSION AC093863.3 GI:18653780
 VERSION
 KEYWORDS
 HTG.
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

misc_feature

1 (bases 1 to 164278)
 Waterston, R.H.
 TITLE The sequence of Homo sapiens clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 164278)
 Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (10-SEP-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 3 (bases 1 to 164278)
 Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (13-FEB-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 On Feb 13, 2002 this sequence version replaced gi:16259166.

misc_feature

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 On Feb 13, 2002 this sequence version replaced gi:16259166.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC

web site: <http://genome.wustl.edu/gsc/index.shtml>
 Contact: submissions@watson.wustl.edu
 ----- Project Information -----
 Center project name: H_NH0552110
 Drafting center: WIBR

FEATURES
 source
 1..164278
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="4"
 /clone="RP11-552110"

BASE COUNT 53573 a 30951 c 30884 g 48870 t
 ORIGIN

Alignment Scores:
 Pred. No.: 187 Length: 164278
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 13.79% Indels: 0
 DB: 9 Gaps: 0

US-09-727-892a-99 (1-58) x AC093863 (1-164278)

Oy 50 TTTTtllleGtYrllleYsglu 57

Db 126031 TACATTATGAGTACATTAAGAA 126054

RESULT 9
 AC090539 171690 bp DNA linear PRI 27-NOV-2001
 LOCUS AC090539 Homo sapiens chromosome 8, clone RP11-759A9, complete sequence.
 DEFINITION AC090539
 AC090539 6 GI:16974167
 VERSION
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 TITLE 1 (bases 1 to 171690)
 JOURNAL Homo sapiens chromosome 8, clone RP11-759A9
 REFERENCE
 AUTHORS 2 (bases 1 to 171690)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
 Batra,N., Bastien,V., Boguslavsky,L., Bouckgalter,B., Brown,A.,
 Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
 Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
 Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Karatas,A., Laroque,K., Lamazares,R., Landers,T.,
 Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,
 McHeaters,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
 McHeaters,R., Meldrim,J., Meneus,L., Mihova,T., McKernan,K.,
 Murphy,T., Naylor,J., Nguyen,C., Nordu,C., Notman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
 Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
 Roy,A., Santos,R., Schauer,S., Schupack,M., Seaman,S., Severy,P.,
 Sougnaz,C., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Testfay,S., Theodore,J.,
 Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A.,
 Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
 Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission
 JOURNAL Submitted (02-MAR-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 3 (bases 1 to 171690)
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Batra,N., Bastien,V., Boguslavsky,L., Bouckgalter,B.,
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,

Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
 Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
 Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gind,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Kamat,A., Karatas,A., Kells,C., Laroque,K.,
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
 Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., McHeaters,R., Meldrim,J.,
 Meneus,L., Mihova,T., Mlenaga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Nordu,C., Notman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
 Seaman,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Testfay,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

COMMENT
 JOURNAL Submitted (27-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Nov 18, 2001 this sequence version replaced gi:16905281.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information -----
 Center project name: L12149
 Center clone name: 759_A_9

FEATURES
 source
 1..171690
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8"
 /clone="RP11-759A9"
 /clone_lib="RPCI-11 Human Male BAC"
 1..327
 /rpt_family="L1ME1"
 224..228
 /note="<30 qual SNGI region"
 398..1316
 /rpt_family="L1ME1"
 1317..1358
 /rpt_family="(TATTA)n"
 1359..2598
 /rpt_family="L1ME1"
 2595..2703
 /rpt_family="L1M4"
 5849..5887
 /rpt_family="AT-rich"
 6098..6158
 /rpt_family="(Tc)n"
 6526..6557
 /rpt_family="AT-rich"
 complement(7153..7329)
 /rpt_family="L1PA7"
 7330..8584
 /rpt_family="L1PA7"
 9681..11495
 /rpt_family="L1PA7"
 complement(11507..11609)
 /rpt_family="L2"
 complement(11900..11954)
 /rpt_family="MIR"
 complement(12437..13225)
 /rpt_family="L1PA2"

```

repeat_region      13224..14814      /rpt_family="L1PA2"
repeat_region      complement(15391..15701)
repeat_region      /rpt_family="Alusp"
repeat_region      15765..16044      /rpt_family="Alusp"
repeat_region      16045..16219      /rpt_family="(GAAA)n"
repeat_region      complement(18876..19178)
repeat_region      /rpt_family="Alusp"
repeat_region      21097..21330      /rpt_family="MIR"
repeat_region      complement(22482..23581)
repeat_region      /rpt_family="L1MC3"
repeat_region      complement(23670..23891)
repeat_region      /rpt_family="L1MC3"
repeat_region      complement(23893..24386)
repeat_region      /rpt_family="MER4D"
repeat_region      complement(24469..24544)
repeat_region      /rpt_family="L1MD3"
repeat_region      complement(24863..24950)
repeat_region      /rpt_family="Charlie8"
repeat_region      25032..25056      /rpt_family="(TTA)n"
repeat_region      complement(25057..26154)
repeat_region      /rpt_family="L1PA4"
repeat_region      complement(26155..26261)
repeat_region      /rpt_family="L1MB2"
repeat_region      complement(26318..26852)
repeat_region      /rpt_family="L1MB3"
repeat_region      complement(26853..27231)
repeat_region      /rpt_family="THEIC"
repeat_region      complement(27232..27441)
repeat_region      /rpt_family="L1MB3"
repeat_region      27442..27464      /rpt_family="(TTG)n"
repeat_region      complement(27465..28494)
repeat_region      /rpt_family="L1MB3"
repeat_region      complement(28497..28895)
repeat_region      /rpt_family="L1MB2"
repeat_region      29022..29050      /rpt_family="AT_rich"
repeat_region      29144..29968      /rpt_family="L1R12"
repeat_region      complement(30288..30579)
repeat_region      /rpt_family="Alusx"
repeat_region      30990..33322      /rpt_family="L1MC4"
repeat_region      33366..34157      /rpt_family="L1MC4"
repeat_region      34183..34920      /rpt_family="L1MC4"
repeat_region      34920..35064      /rpt_family="L1MC4"
repeat_region      /rpt_family="L1MC3"
repeat_region      35203..35397      /rpt_family="Tiger1"
repeat_region      35466..35499      /rpt_family="Tiger1"
repeat_region      complement(35500..35648)
repeat_region      /rpt_family="Alusg/x"
repeat_region      35649..37164      /rpt_family="Tiger1"
repeat_region      37165..37473      /rpt_family="Alusg"
repeat_region      37474..38134      /rpt_family="Tiger1"
repeat_region      complement(38135..38259)
repeat_region      /rpt_family="L1PB2"
repeat_region      38260..38304      /rpt_family="Tiger1"
repeat_region      38324..39073      /rpt_family="L1MC4"
repeat_region      39074..39122

```

```

US-09-727-892a-99 (1-58) x AC090539 (1-171690)
Gy 14 Glut1elysGlyHisPheProHis 21
Db 22914 GAATTAAGGCTCATTTCCCTCAT 22937

RESULT 10
AC098858/c 172419 bp DNA linear PRI 03-JAN-2002
LOCUS
DEFINITION Homo sapiens chromosome 4 clone RP11-149A7, complete sequence.
ACCESSION AC098858 AC011960
VERSION AC098858.3 GI:17737053
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 172419)
REFERENCE
AUTHORS Waterston,R.H.
JOURNAL The sequence of Homo sapiens clone
TITLE Unpublished
2 (bases 1 to 172419)
REFERENCE
AUTHORS Waterston,R.H.
JOURNAL Direct Submission
TITLE Submitted (04-NOV-2001) Genome Sequencing Center, Washington
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
3 (bases 1 to 172419)
REFERENCE
AUTHORS Waterston,R.H.
JOURNAL Direct Submission
TITLE Submitted (14-DEC-2001) Genome Sequencing Center, Washington
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 172419)
REFERENCE
AUTHORS Waterston,R.H.
JOURNAL Direct Submission
TITLE Submitted (03-JAN-2002) Genome Sequencing Center, Washington
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Dec 14, 2001 this sequence version replaced gi:17647053.

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: H_NH0149A07
Drafting center: WMBR
-----
FEATURES
Source location/Qualifiers
1..172419

```



```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-149A7"
BASE COUNT      55963 a 29383 c 30032 g 57041 t
ORIGIN
Alignment Scores:
Pred. No.:      194      Length:      172419
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      13.79%      Indels:      0
DB:              9      Gaps:      0

US-09-727-892a-99 (1-58) x AC098858 (1-172419)

Qy      7 79HValleuLeuTyTCysAspGlu 14
        |||||
        ACAGTCCTCTATATGTGATGAA 9103

RESULT 11
AC068217      178776 bp      DNA      linear      HTG 09-JUN-2000
LOCUS      Homo sapiens chromosome 2 clone RP11-70P23 map 2, WORKING DRAFT
DEFINITION      AC068217
ACCESSION      AC068217
VERSION      AC068217.2 GI:8389495
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 178776)
Birtten,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 2, clone RP11-70P23
Unpublished
2 (bases 1 to 178776)
Birtten,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,R.,
Boguslavsky,L., Bouhagalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lien,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., MCPheeters,R.,
McCarthy,J., Menes,L., Minova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teefaye,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,D., Zimmer,A. and Zody,M.
Direct Submission
Submitted (30-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 9, 2000 this sequence version replaced gi:7671277.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WITR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L10121

```

```

----- Summary Statistics
Center clone name: 70.P.23
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 167970 bases at least Q40
Consensus quality: 173942 bases at least Q30
Consensus quality: 176139 bases at least Q20
Insert size: 18200; agarose-fp
Insert size: 177376; sum-of-contigs
Quality coverage: 4.4 in Q20 bases; agarose-fp
Quality coverage: 4.6 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1
691 790: gap of 100 bp
791 4223: contig of 3433 bp in length
4224 4323: gap of 100 bp
4324 7371: contig of 3048 bp in length
7372 7471: gap of 100 bp
7472 11158: contig of 3687 bp in length
11159 11258: gap of 100 bp
11259 15848: contig of 4590 bp in length
15849 15948: gap of 100 bp
15949 21796: contig of 5848 bp in length
21797 21896: gap of 100 bp
21897 31059: contig of 9163 bp in length
31060 31159: gap of 100 bp
31160 42738: contig of 11579 bp in length
42739 42838: gap of 100 bp
42839 54578: contig of 11740 bp in length
54579 54678: gap of 100 bp
54679 67044: contig of 12366 bp in length
67045 67144: gap of 100 bp
67145 80822: contig of 12678 bp in length
80823 80922: gap of 100 bp
80923 97662: contig of 16740 bp in length
97663 97762: gap of 100 bp
97763 121636: contig of 23874 bp in length
121637 121736: gap of 100 bp
121737 146379: contig of 24643 bp in length
146380 146479: gap of 100 bp
146480 178776: contig of 32297 bp in length.
-----
FEATURES
source
Location/Qualifiers
1..178776
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-70P23"
/clone_lib="RPC1-11 Human Male BAC"
1..690
/clone="assembly-fragment
clone.end:SP6
vector.side:left"
791..4223
/clone="assembly-fragment"
4324..7371
/clone="assembly-fragment"
7472..11158
/clone="assembly-fragment"
11259..15848
/clone="assembly-fragment"
15949..21796
/clone="assembly-fragment"
21897..31059

```

```

misc_feature      /note="assembly_fragment"
                  31160..42738
misc_feature      /note="assembly_fragment"
                  42839..54578
misc_feature      /note="assembly_fragment"
                  54679..67044
misc_feature      /note="assembly_fragment"
                  67145..80822
misc_feature      /note="assembly_fragment"
                  clone_end:T7
                  vector_side:right"
misc_feature      80923..97662
                  /note="assembly_fragment"
misc_feature      97763..121636
                  /note="assembly_fragment"
misc_feature      121737..146379
                  /note="assembly_fragment"
misc_feature      146480..178776
                  /note="assembly_fragment"

BASE COUNT      55891 a 32274 c 32935 g 56275 t 1401 others
ORIGIN
Alignment Scores:
Pred. No.:      200      Length:      178776
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      13.79%      Indels:      0
DB:              2      Gaps:      0

US-09-727-892a-99 (1-58) x AC068217 (1-178776)
QY      50      Tyrllelleglutyrllelsglu 57
Db 149436 TACATTATAGAGTACATTAAGAA 149459

RESULT 12
AC022820/c      180899 bp      DNA      linear      PRI 11-DEC-2001
LOCUS          AC022820      Homo sapiens chromosome 8, clone RP11-46G17, complete sequence.
DEFINITION
AC022820
AC022820.11      GI:17488713
KEYWORDS
HTG.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 180899)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 8, clone RP11-46G17
Unpublished
2 (bases 1 to 180899)
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Domino,M., Doyle,M., Feneslor,J.,
Fitzhugh,W., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howard,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marguis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Menus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody M.
Direct Submission
Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE
3 (bases 1 to 180899)

```

```

AUTHORS
Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P.,
Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S.,
Goyette,M., Graham,W., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Larocque,K., Lamazares,C., Macdonald,P., Major,J.,
Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Major,J.,
McPheeters,R., Meldrim,J., Menus,L., Mihova,T., Mlenge,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,D., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Sougnaz,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody M.
Direct Submission
Submitted (03-JUL-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE
4 (bases 1 to 180899)
AUTHORS
Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,W., Grand-Pierre,N.,
Hagos,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Larocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marguis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Menus,L., Mihova,T., Mlenge,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,D., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody M.

```

```

TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (11-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 11, 2001 this sequence version replaced gi:14589634.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: D6116
Center clone name: 46_G.17
-----
Location/Oualifiers
1..180899
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-46G17"
/clone_lib="RPCT-11 Human Male BAC"
repeat_region complement(1660..2111)

```

```

repeat_region      /rpt_family="L1MA9"
                    2347..2645
repeat_region      /rpt_family="Aluv"
                    2657..2677
repeat_region      /rpt_family="AT_rich"
                    3619..3737
repeat_region      /rpt_family="MIR"
                    4412..4497
repeat_region      /rpt_family="(TGA)n"
                    complement(4753..5574)
repeat_region      /rpt_family="L1MEd"
                    5813..5833
repeat_region      /rpt_family="AT_rich"
                    5912..5953
repeat_region      /rpt_family="(TG)n"
                    5968..6122
repeat_region      /rpt_family="MIR"
                    6826..6862
repeat_region      /rpt_family="AT_rich"
                    6943..6972
repeat_region      /rpt_family="AT_rich"
                    8783..10099
repeat_region      /rpt_family="L2"
                    10113..10844
repeat_region      /rpt_family="L1MDa"
                    10831..11924
repeat_region      /rpt_family="L1MDa"
                    11925..12216
repeat_region      /rpt_family="AluSc"
                    12217..12311
repeat_region      /rpt_family="L1MDa"
                    complement(12312..12862)
repeat_region      /rpt_family="MER41A"
                    complement(12892..13415)
repeat_region      /rpt_family="LTR49"
                    13422..13466
repeat_region      /rpt_family="(TCFA)n"
                    complement(13621..14248)
repeat_region      /rpt_family="LTR8A"
                    14265..14407
repeat_region      /rpt_family="L1MDa"
                    14408..14516
repeat_region      /rpt_family="L2"
                    15099..15474
repeat_region      /rpt_family="HAL1"
                    15685..16398
repeat_region      /rpt_family="HAL1"
                    16406..16877
repeat_region      /rpt_family="MER1A"
                    complement(17174..17462)
repeat_region      /rpt_family="L1MC3"
                    17472..17548
repeat_region      /rpt_family="Tigger3(Golem)"
                    17554..17969
repeat_region      /rpt_family="Tigger3(Golem)"
                    complement(18091..18278)
repeat_region      /rpt_family="MER2"
                    complement(18607..18906)
repeat_region      /rpt_family="AluIo"
                    19220..19517
repeat_region      /rpt_family="AluIo"
                    19557..19630
repeat_region      /rpt_family="MIR"
                    20074..21332
repeat_region      /rpt_family="L1MC/D"
                    21334..21436
repeat_region      /rpt_family="L1MB8"
                    complement(21511..21706)
repeat_region      /rpt_family="L1MC3"
                    complement(21713..21936)
repeat_region      /rpt_family="LTR47B"
                    complement(21972..22094)
repeat_region      /rpt_family="LTR47B"

```

```

repeat_region      complement(22095..22428)
repeat_region      /rpt_family="L1MC4"
                    22446..22524
repeat_region      /rpt_family="L1MC1"
                    complement(22517..23959)
repeat_region      /rpt_family="L1MC4"
                    complement(23960..24269)
repeat_region      /rpt_family="AluSc"
                    complement(24270..24612)
repeat_region      /rpt_family="L1MC4"
                    24760..24857
repeat_region      /rpt_family="MIR"
                    complement(24881..24984)
repeat_region      /rpt_family="L2"
                    complement(25070..25890)
repeat_region      /rpt_family="L1P45"
                    27626..27654
repeat_region      /rpt_family="(TG)n"
                    30881..30914

```

```

Alignment Scores:
Pred. No.:      202
Score:          8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match:    13.79%
DB:             9
Gaps:           0

```

US-09-727-892A-99 (1-58) x AC022820 (1-180899)

Oy 14 Glu1elsglyH1sPheProH1s 21

Db 68545 GAAATMAAGGTCATTTCCTCAT 68522

RESULT 13

AC073125

LOCUS

DEFINITION

AC073125

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

TITLE
JOURNAL

COMMENT

Direct Submission
Submitted (10-MAY-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Mar 7, 2001 this sequence version replaced gi:11245625.

----- Genome Center
Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc>

Contact: sapiens@wustl.wustl.edu

----- Summary Statistics

Center project name: H_NH0636L15

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and
sequencing collaboration between the NHGRI Chromosome 7 Mapping
Project (Eric D. Green, Director), John D. McHesron in the
Department of Genetics (Washington University), and the Washington
University Genome Sequencing Center. For additional information
about the map position of this sequence, see
<http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send
<mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa, K., Moon, P. Y., Zhao, B., Frengen, E.,
Tateno, M., Catanesse, J. J. and de Jong, P. J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong
and coworkers at the Roswell Park Cancer Institute
(<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-806J6. Actual start of
this clone is at base position 1 of RP11-636L15; actual end is at
base position 188804 of RP11-636L15.

Location/Qualifiers

1. 188804

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="7"

/map="7"

/clone="RP11-636L15"

/clone_lib="RPCI-11"

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

/rpt_family="L1"
4551. .4853

/rpt_family="Alu"

4854. .5453

/rpt_family="L1"

5454. .6095

/rpt_family="L1"

6096. .6456

/rpt_family="L1"

6459. .6767

/rpt_family="Alu"

6782. .6960

/rpt_family="L1"

7380. .7442

/rpt_family="AT_rich"

8993. .9374

/rpt_family="MaLR"

9386. .9502

/rpt_family="MER94"

10524. .10545

/rpt_family="AT_rich"

10606. .10693

/rpt_family="AT_rich"

10795. .10820

/rpt_family="AT_rich"

10821. .11129

/rpt_family="Alu"

11130. .11156

/rpt_family="AT_rich"

11157. .11602

/rpt_family="L1"

12391. .12481

/rpt_family="GA_rich"

12865. .13165

/rpt_family="Alu"

14122. .14466

/rpt_family="AchoBo"

16363. .16615

/rpt_family="Alu"

20707. .20755

/rpt_family="L1"

21115. .21339

/rpt_family="Alu"

22187. .22226

/rpt_family="AT_rich"

22255. .22284

/rpt_family="GGAA)n"

22625. .22923

/rpt_family="Alu"

23305. .23385

/rpt_family="L2"

23548. .23936

/rpt_family="L1"

23937. .24385

/rpt_family="ERV.L"

24386. .24505

/rpt_family="L1"

24570. .24609

/rpt_family="GGAA)n"

24667. .24730

/rpt_family="L1"

25022. .25331

/rpt_family="L1"

25332. .25363

/rpt_family="TGAA)n"

25364. .25563

/rpt_family="L1"

25932. .25960

/rpt_family="TC)n"

26567. .27821

/rpt_family="L1"

28958. .29060

/rpt_family="MIR"

```

repeat_region 29328..29379
/rpt_family="(CATATA)n"
repeat_region 29565..29605
/rpt_family="AT_rich"
repeat_region 29917..29996
/rpt_family="L2"
repeat_region 30288..30407
/rpt_family="L1"
repeat_region 30409..30471
/rpt_family="MALR"
repeat_region 30472..30932
/rpt_family="ERV1"
repeat_region 30933..31276
/rpt_family="MALR"
repeat_region 31277..32895
/rpt_family="MALR"
repeat_region 32900..33098
/rpt_family="MALR"

Alignment Scores:
Pred. No.: 209 Length: 188804
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.79% Indels: 0
DB: 9 Gaps: 0

US-09-727-892a-99 (1-58) x AC073125 (1-188804)
QY 18 HispPProHISGlnIleSerMet 25
Db 107639 CATTTTCACATCAATATCTATG 107662
|||||
RESULT 14
AC024422
LOCUS AC024422 194215 bp DNA linear HTG 16-MAR-2000
DEFINITION Homo sapiens chromosome 18 clone RP11-130C2 map 18, WORKING DRAFT
SEQUENCE 19 unordered pieces.
AC024422
AC024422.2 GI:7249325
HTG: HTGS_PHASE1 HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 194215)
Birren,B., Linton,L., Nussbaum,C. and Lander,E.
Homo sapiens chromosome 18, clone RP11-130C2
Unpublished
2 (bases 1 to 194215)
Birren,B., Linton,L., Nussbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavsky,L.,
Bouhagelter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Chepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeRellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M.,
Fenster,J., Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D.,
Galagan,J., Gardyna,S., Glade,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karakas,A.,
Klein,J., Landers,T., Largocque,K., Lehoczy,J., Levine,R.,
Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N., McCarthy,M.,
McWan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrum,J.,
Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M.,
Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Tirelli,A.,
Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
Wu,X., Wyman,D., Ye,W.J., Young,G., Zahoun,J., Zimmer,A. and
Zody,M.
Direct Submission
Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

```

COMMENT

On Mar 16, 2000 this sequence version replaced gi:7108218.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L6917

Center clone name: 130_C_2

Summary Statistics

Sequencing vector: M13: M77815: 100% of reads

Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly program: Phrap: version 0.960731

Consensus quality: 180294 bases at least Q40

Consensus quality: 187122 bases at least Q30

Consensus quality: 190103 bases at least Q20

Insert size: 192415: sum-of-contigs

Quality coverage: 4.0 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
 consists of 19 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

```

1 1752: contig of 1752 bp in length
1753 1852: gap of 100 bp
1853 2348: contig of 496 bp in length
2349 2448: gap of 100 bp
2449 4672: contig of 2224 bp in length
4673 4772: gap of 100 bp
4773 7822: contig of 3050 bp in length
7823 7922: gap of 100 bp
7923 10519: contig of 2597 bp in length
10520 10619: gap of 100 bp
10620 14451: contig of 3832 bp in length
14452 14551: gap of 100 bp
14552 18402: contig of 3851 bp in length
18403 18502: gap of 100 bp
18503 21919: contig of 3417 bp in length
21920 22019: gap of 100 bp
22020 29789: contig of 7770 bp in length
29790 29889: gap of 100 bp
29890 37699: contig of 7810 bp in length
37700 37799: gap of 100 bp
37800 50710: contig of 12911 bp in length
50711 50810: gap of 100 bp
50811 62984: contig of 12174 bp in length
62985 63084: gap of 100 bp
63085 74715: contig of 11631 bp in length
74716 74815: gap of 100 bp
74816 88792: contig of 13977 bp in length
88793 88892: gap of 100 bp
88893 102951: contig of 14059 bp in length
102952 103051: gap of 100 bp
103052 119776: contig of 16725 bp in length
119777 119876: gap of 100 bp
119877 135778: contig of 15902 bp in length
135779 135878: gap of 100 bp
135879 154088: contig of 18210 bp in length
154089 154188: gap of 100 bp
154189 194215: contig of 40027 bp in length.

```

FEATURES

SOURCE

1..194215
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="18"
 /map="18"

JOURNAL Submitted (03-AUG-1999) Molecular and Human Genetics/HMT, Baylor College of Medicine, T634, Mail stop BCM235, One Baylor Plaza, Houston, TX 77030, USA

FEATURES
SOURCE
Location/Qualifiers
1. 469

/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="3"
/map="85E5-11"
/note="flanking the 3'-end of the dmt(S048103) P element;
dmt(S048103) P element is inserted in the dalmatian (dmt)
gene
Sequence recovery method: 7.0 kb A16 genomic fragment was
isolated by plasmid rescue with EcoRI restriction enzyme;
fragment was directly sequenced at the proximal end
(relative to the site of P element insertion)"

BASE COUNT 147 a 117 c 113 g 92 t
ORIGIN

Alignment Scores:

Pred. No.: 25.9 length: 469
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 3 Gaps: 0

US-09-727-892a-99 (1-58) x AF174682 (1-469)

QY 40 GlutyrAsnLeupherLys 46
|||||
Db 45 GAATTAATTTGTTTACCATA 25

RESULT 17

AX284699/c AX284699 509 bp DNA linear PAT 20-NOV-2001
DEFINITION Sequence 504 from Patent WO0179556.
ACCESSION AX284699
VERSION AX284699.1 GI:17045387
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (sites) Brown, J.L., Bolt, A. and van Hufel, C.
Novel genes, compositions, and methods for the identification,
assessment, prevention, and therapy of human cancers
Patent: WO 0179556-A 504 25-OCT-2001;
Millennium Predictive Medicine, Inc. (US)

FEATURES
SOURCE
Location/Qualifiers
1. 509
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 150 a 97 c 101 g 161 t
ORIGIN

Alignment Scores:

Pred. No.: 27.6 length: 509
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 6 Gaps: 0

US-09-727-892a-99 (1-58) x AX284699 (1-509)

QY 6 LysThrValLeuTyrCys 12
|||||

Db 486 AAAACAGCTCTGCTGATTGC 466

RESULT 18
G58636/c

LOCUS G58636 573 bp DNA linear STS 30-MAR-2000
DEFINITION SHGC-104707 Human Homo sapiens STS genomic, sequence tagged site.
ACCESSION G58636
VERSION G58636.1 GI:6123805
KEYWORDS STS.

SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 573)
Olivier, M. and Cox, D.R.
Unpublished, Olivier, M., Cox, D.R. (2000)
unpublished

JOURNAL

Contact: Michael Olivier, David R. Cox
Stanford Human Genome Center
Stanford University School of Medicine
4005 Miranda Ave., 2nd Fl., Palo Alto, CA 94025, USA
Tel: (650) 320-5800
Fax: (650) 320-5801
Email: Olivier@shgc.stanford.edu
Primer A: AAAGCAGCTGCGAGAAAGTG
Primer B: TAAGAGCCACATCCACAGAAAT
STS size: 296
PCR Profile:

Initial incubation: 95 degrees C for 10 minutes
Denaturation: 94 degrees C for 30 seconds
Annealing: 60 degrees C for 30 seconds
Polymerization: 72 degrees C for 23 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9700
Protocol:

Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Amplifrag Gold Polymerase: 0.07 units/ul
Total Vol: 5 ul

Buffer: MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3

BAC ends sequenced at TIGR from the RPC11 BAC library. Designed
and developed at the Stanford Human Genome Center.

FEATURES
SOURCE
Location/Qualifiers
1. 573
/organism="Homo sapiens"
/db_xref="taxon:9606"

STS
primer_bind
158..453
complement(431..453)

BASE COUNT 149 a 120 c 125 g 177 t 2 others
ORIGIN

Alignment Scores:

Pred. No.: 30.4 length: 573
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 11 Gaps: 0

US-09-727-892a-99 (1-58) x G58636 (1-573)

QY 3 ArgLysTyrLysThrValLeu 9
|||||

Db 85 AGAAATATCAAAACTGTGTTA 65

RESULT 19

BAU10498/c 673 bp DNA linear BCT 11-OCT-1995
 LOCUS BAU10498
 DEFINITION Buchnera aphidicola from Rhopalosiphum padi shikimate dehydrogenase (aroF) gene, partial cds, tRNA-Glu gene, complete sequence, and 23S rRNA gene, partial sequence.
 ACCESSION U10498.1 GI:854715
 VERSION
 KEYWORDS Buchnera aphidicola.
 SOURCE Buchnera aphidicola
 ORGANISM Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 REFERENCE 1 (bases 1 to 673)
 AUTHORS Roubhakhsh, D. and Baumann, P.
 TITLE Characterization of a putative 23S-5S rRNA operon of Buchnera aphidicola (endosymbiont of aphids) unlinked to the 16S rRNA-encoding gene
 JOURNAL Gene 155 (1), 107-112 (1995)
 MEDLINE 95212914
 REFERENCE 2 (bases 1 to 673)
 AUTHORS Roubhakhsh, D.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1994) Dadbeh Roubhakhsh, Microbiology, University of California at Davis, Davis, CA 95616, USA
 FEATURES
 source
 1. .673
 /organism="Buchnera aphidicola"
 /specific_host="Rhopalosiphum padi"
 /db_xref="taxon:9"
 1. .69
 /gene="aroF"
 <1. .69
 /gene="aroE"
 /transl_start=1
 /transl_table=11
 /product="shikimate dehydrogenase"
 /protein_id="AA09127.1"
 /db_xref="GI:854716"
 /translation="SFLEWHVPEETNYVDALSN"
 145. .150
 165. .170
 229. .302
 /product="tRNA-Glu"
 /anticodon="(pos:238. .240, aa:Glu)
 464. .>673
 /product="23S ribosomal RNA"
 BASE COUNT 241 a 89 c 107 g 236 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 34.6 Length: 673
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 DB: 1 Gaps: 0
 US-09-727-892a-99 (1-58) x BAU10498 (1-673)
 QY 50 TyrIleIleGluTyrIleYs 56
 ||||||||||||||||||||
 Db 132 TATATATAGATATATATAA 112
 RESULT 20
 D89182/c 833 bp mRNA linear PLN 13-MAR-1998
 LOCUS D89182
 DEFINITION Schizosaccharomyces pombe mRNA, partial cds, clone: SY 1014.
 ACCESSION D89182
 VERSION D89182.1 GI:1749571
 KEYWORDS Schizosaccharomyces pombe (strain:PR745) cDNA to mRNA, clone_lib:library of H. Nojima clone:ST1014.
 SOURCE Schizosaccharomyces pombe
 ORGANISM Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

Schizosaccharomycetales; Schizosaccharomycetaceae;
 Schizosaccharomycetes.
 REFERENCE 1 (bases 1 to 833)
 AUTHORS Yoshioaka, S.
 TITLE Direct Submission
 JOURNAL Submitted (15-NOV-1996) Sachio Yoshioaka, Tsukita Cell Axis Project
 ERATO-JST, Kyoto Research Park, 17 Chidouji Minamimachi,
 Shimokyo-ku, Kyoto, Kyoto 600, Japan
 (E-mail: syoshi@cell.tsukita.jst.go.jp, Tel: +81-75-315-7913,
 Fax: +81-75-315-6420)
 2 (sites)
 REFERENCE Yoshioaka, S., Kato, K., Nakai, K., Okayama, H. and Nojima, H.
 AUTHORS Identification of open reading frames in Schizosaccharomycetes pombe
 TITLE CDNAS
 JOURNAL DNA Res. 4 (6), 363-369 (1997)
 MEDLINE 98162722
 REFERENCE
 source
 1. .833
 /organism="Schizosaccharomyces pombe"
 /strain="PR745"
 /db_xref="taxon:4896"
 /clone="SY1014"
 /clone_lib="library of H. Nojima"
 232. .594
 /note="Similar to Saccharomyces cerevisiae Ipg10p,
 SWISS-PROT Accession Number U43281"
 /codon_start=1
 /protein_id="BA13844.1"
 /db_xref="GI:1749572"
 /translation="MOSMOPSTYDRKMKAMGSAAGLGIFGCVAVLRGPRPG
 FURTLCQYMTSATRFGFFPMSISVIRNEDPLIQSGSHMQRLLNENANSSRIFLA
 AMQQAQSPKRSNEVAEC"
 BASE COUNT 262 a 149 c 133 g 289 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 41 Length: 833
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 DB: 8 Gaps: 0
 US-09-727-892a-99 (1-58) x D89182 (1-833)
 QY 33 LysValValTyrSerTyrTyr 39
 ||||||||||||||||||||
 Db 152 AAGGTAGTATACGTATATAT 132
 RESULT 21
 GIU94701 909 bp DNA linear INV 19-MAR-1998
 LOCUS GIU94701
 DEFINITION Giardia intestinalis vacuolar ATPase proteolipid subunit (vma-3)
 ACCESSION U94701
 VERSION U94701.1 GI:2978500
 KEYWORDS Giardia intestinalis.
 SOURCE Giardia intestinalis
 ORGANISM Eukaryota; Diplomonadida; Hexamitidae; Giardinales; Giardia.
 REFERENCE 1 (bases 1 to 909)
 AUTHORS Peattie, D.A., Alonso, R.A., Hahn, A. and Caulfield, J.P.
 TITLE Ultrastructural localization of giardins to the edges of disk microribbons of Giardia lamblia and the nucleotide and deduced protein sequence of alpha giardin
 JOURNAL J. Cell Biol. 109 (5), 2323-2335 (1988)
 MEDLINE 90037227
 REFERENCE 2 (bases 1 to 909)
 AUTHORS Hilaris, E. and Gogarten, J.P.
 TITLE The vacuolar proteolipid of Giardia: evolution of V/F/A-ATPase function reflections on the prokaryote-eukaryote transition
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 909)

AUTHORS Hilario,E. and Gogarten,J.P.
 TITLE Direct Submission
 JOURNAL Submitted (19-MAR-1997) Molecular and Cell Biology, University of Connecticut, 75 N. Eagleville Rd. U-44, Storrs, CT 06269, USA

FEATURES

location/Qualifiers
 1..909

organism="Giardia intestinalis"
 strain="WB C6"
 /db_xref="taxon:5741"
 /clone_lib="trophozoite genomic DNA library in lambda Zap II kindly provided by Dr. Frances D. Gillin, University of California at San Diego"
 /dev_stage="trophozoite"
 309..842
 /gene="vma-3"
 309..842
 /gene="vma-3"
 /EC_number="3.6.1.35"
 /function="involved in proton translocation"
 /note="V-ATPase subunit c; V-ATPase proteolipid; vacuolar proton pumping adenosine triphosphatase proteolipid subunit"
 /codon_start=1
 /product="vacuolar ATPase proteolipid subunit"
 /protein_id="AAC06133.1"
 /db_xref="GI:2978501"
 /translation="MSSIDSPVAEKCPAGSFWMLGOVAVVSSIGAAVGTAKAGGAGCCGLINAPYTKTLPIYAGISITVLTSLINSRVSTYTNMPYVSTAH LSTRSADSNFCVNNVNO"
 838..843
 /note="putative site proposed for Giardia genes"
 /citation=[1]

polyA_signal

BASE COUNT 201 a 217 c 235 g 256 t

ORIGIN

Alignment Scores:

Pred. No.: 44 Length: 909
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 DB: 3 Gaps: 0

US-09-727-892a-99 (1-58) x GIU94701 (1-909)

OY 7 ThrValLeuLeuTyrCysAsp 13
 |||||

Db 141 ACCGTTCTCTCTACTCGCAT 161

RESULT 22

CNS06J7/c

LOCUS CNS06J7 948 bp DNA linear STS 10-JAN-2001
 DEFINITION T7 end of clone AS0AA028A10 of library AS0AA from strain CLIB 533
 of Saccharomyces bayanus, sequence tagged site.

ACCESSION ALA01477

VERSION ALA01477.1 GI:12159149

KEYWORDS

SOURCE

ORGANISM

Saccharomyces bayanus.
 Saccharomyces bayanus.
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

2 (bases 1 to 948)
 Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
 Boitlin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
 deMontigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,

Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
 Saurin,M., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
 Wincker,P. and Weissenbach,J.
 Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of
 yeast species for molecular evolution studies(1)

FEBS Lett. 487 (1), 3-12 (2000)

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

CDS	complement(339. .1586)
-----	------------------------

Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Kim, C.J., Chen, H., Cheuk, R., Koesema, E., Meyers, M.C., Shin, P., Tracy, S.E., Ban, J., Bowser, L., Chung, M.K., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yu, G., Davis, R.W., Theologis, A., and Ecker, J.R.

Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source Location/Qualifiers

1..1854
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAFL04-20-P06(R12636)"
/note="ecotype: Columbia"
1..288
289..1503
/note="putative AVR elicitor response protein"
/codon_start=-1
/product="At1g05170/YUPB12_22"
/protein_id="AAK63859.1"
/db_xref="GI:14488078"
/translation="MSAKIKGEYSRSFVSRRKTTILCLGFCVGFETNRMNIPES
KGMSPVTEAEERLKIYSEGNPKAKVEKRPOLDFEVANTHIALDTLTKTSLMS
ELAARAVQESLNGAPLSDPMKKQPOEORFLMYVGINATSESRKRRDSIRATWMP
QGERKKLEERKGIIRFVIGHSATGILDRALAEADRKGDRLDHEGYLELSG
KTRTFSTASMDADYVVDVVDVHNATLIGETLVAHRKPRVYIGCMKSGVLSQ
KGVYHEPEYKGEKNGNKYFRHATGQLYAISRLASYISINCHVLRHYANEDVLSG
WFIGIDVKAHIDRRLCCGTPPCPCWKAQACNICVASFDMSCSGICRSADRKEVHRRC
GEGERKALMSATP"

3'UTR
BASE COUNT 547 a 324 c 449 g 534 t
ORIGIN 1504..1854

Alignment Scores:

Pred. No.: 77.9 Length: 1854
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 8 Gaps: 0

US-09-727-892a-99 (1-58) x AF389287 (1-1854)

QY 5 TyrlYThrValLeuLeuTyr 11

Db 629 TACAAACGGTCCTTTAT 649

RESULT 28

YSCA49A 1877 bp DNA linear PLN 27-APR-1993

LOCUS YSCA49A 1877 bp DNA linear PLN 27-APR-1993

DEFINITION S.cerevisiae RNA polymerase alpha-subunit (RPA49) gene, complete

cds. M96600.1 GI:170955

ACCESSION RNA polymerase alpha-subunit.

KEYWORDS Saccharomyces cerevisiae DNA.

SOURCE Saccharomyces cerevisiae

ORGANISM Eukaryota; Fungi; Ascomycota; Saccharomycetaceae; Saccharomycetes;

REFERENCE 1 (bases 1 to 1877)

AUTHORS Liljelund, P., Mariotte, S., Buhler, J.-M. and Sentenac, A.

TITLE Characterization and mutagenesis of the gene encoding A49 subunit

JOURNAL of RNA polymerase A Saccharomyces cerevisiae

MEDLINE Proc. Natl. Acad. Sci. U.S.A. 89, 9302-9305 (1992)

FEATURES 93028452 Location/Qualifiers

source I..1877

/organism="Saccharomyces cerevisiae"
/db_xref="taxon:4932"
362..1609
/gene="RPA49"
362..1609
/gene="RPA49"

BASE COUNT 611 a 346 c 360 g 560 t
ORIGIN

Alignment Scores:

Pred. No.: 78.7 Length: 1877
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 8 Gaps: 0

US-09-727-892a-99 (1-58) x YSCA49A (1-1877)

QY 4 LysTyrlYThrValLeuLeu 10

Db 1685 AAATATAGCGTCCTCTG 1705

RESULT 29

AF311953/c 1954 bp DNA linear PLN 02-NOV-2001

LOCUS AF311953 1954 bp DNA linear PLN 02-NOV-2001

DEFINITION Lycopersicon esculentum expansin (Exp8) gene, partial cds.

ACCESSION AF311953

VERSION AF311953.1 GI:16588450

KEYWORDS tomato.

SOURCE Lycopersicon esculentum

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE 1 (bases 1 to 1954)

AUTHORS Chen, F. and Bradford, K.J.

TITLE Using germination-specific promoter in the construction of

JOURNAL nondestructive reporter for single-seed germination assay

MEDLINE Submitted (09-OCT-2000) Vegetable Crops, University of California,

FEATURES 1954 Location/Qualifiers

source 1..1954

gene /organism="Lycopersicon esculentum"

misc_feature /db_xref="taxon:4081"

1..1954

/gene="Exp8"

/gene="Exp8"

/note="contains promoter and 5' UTR"

/gene="Exp8"

/product="expansin"

1952..>1954

/gene="Exp8"

/codon_start=-1

```
/product="expansin"
/protein_id="AAL26805.1"
/db_xref="GI:16588451"
/translation="M"

BASE COUNT      727 a      263 c      276 g      688 t
ORIGIN

Alignment Scores:
Pred. No.:      81.3      Length:      1954
Score:          7.00      Matches:      7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     12.07%      Indels:      0
DB:              8      Gaps:          0

US-09-727-892A-99 (1-58) x AF311953 (1-1954)

QY      20 PROHISGLNIESERMETPHE 26
        |||||||
Db       129 CCACATCAGATATCGATGTTT 109

RESULT 30
HSU42303      1994 bp      mRNA      linear      PRI 04-APR-1996
DEFINITION   Human ATP:D-hexose 6-phosphotransferase mRNA, partial cds.
ACCESSION    U42303
VERSION      U42303.1      GI:1147780
KEYWORDS
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 1994)
AUTHORS      Palma, F., Agostini, D., Mason, P., Daccha, M., Piccoli, G.,
              Biagiarelli, B., Fiorani, M. and Stocchi, V.
              Purification and characterization of the carboxyl-domain of human
              hexokinase type III expressed as fusion protein
              Mol. Cell. Biochem. 155 (1), 23-29 (1996)
96351387
JOURNAL      2 (bases 1 to 1994)
MEDLINE      Palma, F.
REFERENCE    Direct Submission
TITLE        Submitted (06-DEC-1995) Francesco Palma, Istituto di Chimica
              Biologica 'Giorgio Fornaini', Via Saffi, Urbino, 2 - 61029, Italy
AUTHORS
JOURNAL
FEATURES
SOURCE
1. 1994
   location/Qualifiers
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone="4.21"
   /tissue_type="liver"
   /clone_lib="CLONTECH HL115a"
1. 82
<83...1783
/EC_number="2.7.1.1"
/function="D-glucose phosphorylation"
/note="hexokinase type III"
/codon_start=1
/product="ATP:D-hexose 6-phosphotransferase"
/protein_id="AAC50422.1"
/db_xref="GI:1147781"
/translation="PSTGAARVAHAILODLGSPASDELVEQVHCAVCAVCTRAALCAA
ALAAVLSCLOHSREOQTLIVAVATGRCOHPFRCILQGVLLAPECVSLIPSY
DGGRGVAMTAVANARLAHRRLLETLAPRLHDDLAANOQMRKAMRGSGSES
SSRMLPTFEVATLTGSRGDLADLGSTRVULVYTTGVQVITSEIYSIPETVAAG
SABGLFDHYDCIVDFQOGKLSGCSLPFGTFSFPCRGLDGLILNWKGRKAD
CEGDVVHSLREALTRQAAVELNVAIVNDFVTGMSCGYEDPCEIGLIVGTQNA
YMEELRNVAAGVPGDSGHMCINMEGALDGSILMRTREPASVDQASINQKRFK
IISGMVGEIYRHLILHTLSGLVLFRCQOIORLOTROIPTKFLSEISSLAROYR
ALIEDGLPLTSDALAVLEVCQVAVSRPOLQDAGAAVAEKKRERRGLEELAVSYG
VDGLVTKLHPRFSSLVAAIVRELAPRCVVFVLOSDESGKALAVTVACHLTLLTRY
"

BASE COUNT      355 a      590 c      631 g      418 t
ORIGIN
```

```
Alignment Scores:
Pred. No.:      82.6      Length:      1994
Score:          7.00      Matches:      7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     12.07%      Indels:      0
DB:              9      Gaps:          0

US-09-727-892A-99 (1-58) x HSU42303 (1-1994)

QY      41 TYRASNLEUPHERHRLYSLYS 47
        |||||||
Db       1969 TATAATTATTTTACAAAAAAA 1989

RESULT 31
SCYNL249C/c      2012 bp      DNA      linear      PLN 11-AUG-1997
LOCUS        SCYNL249C
DEFINITION   S.cerevisiae chromosome XIV reading frame ORF YNL249c.
ACCESSION    Z71525 Y13139
VERSION      Z71525.1      GI:1302290
KEYWORDS
SOURCE       Baker's yeast.
ORGANISM     Saccharomyces cerevisiae
              Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
              Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE    1 (bases 1 to 2012)
AUTHORS      Sen-Gupta, M., Gueldeener, U., Beinhauer, J., Fiedler, T. and
              Hegemann, J.H.
              Unpublished
JOURNAL      2 (bases 1 to 2012)
MEDLINE
REFERENCE    Submitted (29-APR-1996) Data collected by MIPS on behalf of the
              European yeast chromosome XIV sequencing project. MIPS at the
              Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a D-82152
              Martinsried, FRG; E-mail: Mewes@mips.emblnet.org
AUTHORS
JOURNAL
FEATURES
SOURCE
1. 2012
   location/Qualifiers
   /organism="Saccharomyces cerevisiae"
   /db_xref="taxon:4932"
   /chromosome="XIV"
   /complement(46.1674)
   /gene="MPA43"
   /complement(46.1674)
   /gene="MPA43"
   /note="ORF YNL249c"
   /codon_start=1
   /protein_id="CAA96156.1"
   /db_xref="GI:1302291"
   /db_xref="SWISS-PROT:P53583"
   /translation="MNSSLQYGVGIDYDSSSARIGVNYNYNDALLMAQEPVYQDS
SKRSKRFWQKSTELIKALQKLCLOKLTNREVSKGVSATSLAIFEDRTSNMLIPY
PNEVDNVLFMDSSAVNBOCMQCPDOLLDYLGKVPKGVKLVYFDYSHLRD
KHEHIFDLHOYIAVELSRLEYMNEIEGLGRENLDGNGDGEVSGMSSFYKNIINLPS
NVSIGTSLVANKHISTTVVASCIDSVASMPVAVSPHLETSLFMIAGSSCYVGGTII
SDPRIPGVAGPGEIILNRDGEVSYAAGSCTGKLTIELHPESHCAKRIIKDGDYIO
VLEGTITDIENKGLSHILTKRDFETFGDGNTPPADPRIKSFTIGESTIDTSMNL
TYKTCILIEFLSFTKLIIDTFQENENSHIHIEKLRIISGQAKNERLLSLIVNNGVA
IIRKENVDMMGIGAVYLAASAKERKOLADVITERDISNDEKFEISLAERYLQNDISI
LRLKLCVKAHHIDMAKQOKRYKRLVDEVOHL"

BASE COUNT      633 a      387 c      323 g      669 t
ORIGIN

Alignment Scores:
Pred. No.:      83.2      Length:      2012
Score:          7.00      Matches:      7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     12.07%      Indels:      0
DB:              8      Gaps:          0

US-09-727-892A-99 (1-58) x SCYNL249C (1-2012)
```

Qy 4 Lysyltyrosylthreylleuleu 10
Db 1936 AATATAAGAGGTCCTCTG 1916

RESULT 32
AF131831 2067 bp mRNA linear PRI 12-MAR-1999
LOCUS Homo sapiens clone 25186 mRNA sequence.
DEFINITION AF131831
ACCESSION AF131831.1 GI:4406669
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 2067)
TITLE Anderson, B., Wentland, M.A., Ricafrente, J.Y., Liu, W. and Gibbs, R.A.
JOURNAL A 'double adaptor' method for improved shotgun library construction
MEDLINE Anal. Biochem. 236 (1), 107-113 (1996)
REFERENCE 2 (bases 1 to 2067)
AUTHORS Yu, W., Anderson, B., Morley, K.C., Muzny, D.M., Ding, Y., Liu, W.,
TITLE Ricafrente, J.Y., Wentland, M.A., Lennon, G. and Gibbs, R.A.
JOURNAL Large-scale concatenation cDNA sequencing
MEDLINE Genome Res. 7 (4), 353-358 (1997)
AUTHORS Mei, G., Yu, W. and Gibbs, R.A.
TITLE Direct Submission
JOURNAL Submitted (26-FEB-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza Rm N1521, Houston, TX 77030, USA
REMARK Contact gme1@bcm.tmc.edu for more information.
FEATURES
Source
1..2067
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="25186"
/note="This clone is similar to human chromosome 5 PAC
clone 170m10 with GenBank Accession Number AC004622. The
I.M.A.G.E. Consortium clone ID number is 25186 and the
library (INIB) was derived from human female infant brain
tissue."

BASE COUNT 588 a 370 c 384 g 725 t
ORIGIN

Alignment Scores:
Pred. No.: 85 Length: 2067
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 9 Gaps: 0

US-09-727-892a-99 (1-58) x AF131831 (1-2067)

Qy 46 Lysyltyrosylthreylleuleu 52
Db 1950 AAAAAGTATGCTTATATTATA 1930

RESULT 33
HS085048/c 2163 bp mRNA linear PRI 16-JAN-1998
LOCUS Homo sapiens cyclic AMP specific phosphodiesterase mRNA, partial
DEFINITION cds.
ACCESSION U85048
VERSION U85048.1 GI:2580521
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 2163)
TITLE Huston, E., Lumb, S., Russell, A., Catterall, C., Ross, A.H.,
Steele, M.R., Bolger, G.B., Perry, M.J., Owens, R.J. and Houslay, M.D.
JOURNAL Molecular cloning and transient expression in COS7 cells of a novel
MEDLINE human PDE4B cAMP-specific phosphodiesterase, HSPDE4B3
REFERENCE 98041898
AUTHORS 2 (bases 1 to 2163)
TITLE Huston, E., Lumb, S., Russell, A., Catterall, C., Perry, M.J.,
Owens, R.J. and Houslay, M.D.
JOURNAL Direct Submission
SUBMITTED (13-JAN-1997) Biochemistry & Molecular Biology, IBIS,
University of Glasgow, Wolfson Building, University Avenue,
Glasgow, Scotland G12 8QQ, UK
FEATURES
Source
1..2163
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
1..>2163
/note="splice variant PDE4B3"
/codon_start=1
/product="cyclic AMP specific phosphodiesterase"
/protein_id="AAB96381.1"
/db_xref="GI:2580522"
/translation="MTAKDSKELTASEPEVCIKFKEQMHLLELPRLPGNRPSPK
ISPSRSPNSPCFFKRLVYKSIQRARRFVATCPDVENGPSRSPDPQASSAG
LVLPATPGHSGORRESPIYRSDSDYDLSPKAMRSNLSRBOHGDLYTPAYLAS
LRYRNRFITLTLNLTGTSNKRSPAASQPPSYRNPRPESYCKLAETLELDWCIDOL
ETIQTRVSEMAKSKFRMLNDELTHLSMSKSGVSEYISNTPFLKQNDVEIPSP
TQDREKKKKQOQMLTQISGVKIKMHSLSNNYSISFGVNTNEDHLAKELHDKWG
LNFENAGVSHNPLICIMYALFOERDLTKFRISDFEITYMTMLDEPHSDVAYHN
SLAAADVAQSTHYLSTPALDAVFTLELITAAIFAAIHDVHPGVSNQFLINTSEL
ALWNVSEVLENHLLAVGFKILOEHCIDIFMNTKROTLERKAVIDVATYADMSKM
SLADAKTWETKKTYSQVLLDNTDTIOVLRMVNCADLSNPKSLLEYLQWQTR
IMEEFQOQKEREGEIMETSPQCDKFTASVSEYGFIDITYHPLMETRADLVQDQ
DIDDTLEDNRMVQSMIPQSPPLDEQNRDCQGLMEKQFELTLDEEDSREPERGE
GHSFSTKTLCTVIDENRSLGETDIDATEDKSPVDY"

BASE COUNT 657 a 514 c 482 g 510 t
ORIGIN

Alignment Scores:
Pred. No.: 88.2 Length: 2163
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 9 Gaps: 0

US-09-727-892a-99 (1-58) x HS085048 (1-2163)

Qy 16 Lysylglycylthreylleuleu 22
Db 259 AAGGCGCATTTCCACATCAA 239

RESULT 34
AF326555/c 2214 bp mRNA linear ROD 02-DEC-2001
LOCUS Mus musculus C57BL/6J phosphodiesterase 4B (Pde4b) mRNA, complete
DEFINITION cds.
ACCESSION AF326555
VERSION AF326555.1 GI:17225436
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS 1 (bases 1 to 2214)
TITLE Fehr, C., Belknap, J.K., Crabbe, J.C. and Buck, K.J.
JOURNAL High resolution mapping of a quantitative trait locus for acute
ethanol withdrawal on mouse chromosome 4 and characterization of
potential candidate genes
Unpublished

REFERENCE 2 (bases 1 to 2214)
AUTHORS Fehr, C., Belknap, J. K., Crabbe, J. C. and Buck, K. J.
TITLE Direct Submission
JOURNAL Submitted (06-DEC-2000) Department of Behavioral Neuroscience,
Portland Alcohol Research Center, Oregon Health Sciences University
and Veterans Affairs Medical Center Portland (Research Services),
3710 SW US Veterans Hospital Road, Portland, OR 97201, USA

FEATURES
source
1..2214
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/chromosome="4"
/map="46.8 cm"
/tissue_type="whole brain"
/dev_stage="adult"
1..2214
/gene="Pde4b"
23..2188
/note="cAMP specific"
/codon_start=1
/product="phosphodiesterase 4B"
/protein_id="AAL37401.1"
/db_xref="GI:17225437"
/translation="MTAKNSPKETASSESEVCITKEKQMRLELELPKLGQRPTSPK
ISPRSPRNSPCFFRKLLVNSITQRRTFYAHTCFVENGSPSGRSLPDPQAGSSSG
LVLHAAPPGHSQRESFLYRSDSDYDLSPKAMSNSSLPSQHGDDLIIVPPAVALS
LRVARNFTLLTNLHGAPNKRSPASQAPVRSVLSQESYOKLAMELELDLCLDL
ETIOTYRSVSEMANFKPKRMNRELTHSEMSRGNOVEYISNPLDKONDVIPS
TOKRREKKKOOLMTQISGVKLLHSSSLNNTSISRGVMTENEDHLAKLEEDLNKNG
LINFVAGYSHNRPLCTIMYAIROERDLTKFKISSPTPTMYMTLEDHHSVAYAN
SLRADVAOSTHYLSTPALDAVFTDLEILAIFPAALIHVDHGVSNQPLINTSEL
ALMINDSVLENHLLAVGFKLQEHCDIFQNLTKQKQRLRKWIDMVLATDMSKIM
SLADLKTWEETKKTSSGVLLDNYDRIOVLNMYCADLSNPTKSLLEYRQWTD
IMEEFFQGGKERERGMETSPMCDKHTASVEKSGVGFIDYIVHPLMETWADLVQPDQ
DIIDLTLEDNRNWQSMIPQSPPLDERSRDCGLMEKFOPELLEBEDSGPERKEGE
GHSYFSSTKLCVDPENRDSLEETDIDINTEKSPIDT"

BASE COUNT 645 a 552 c 527 g 490 t

ORIGIN

Alignment Scores:
Pred. No.: 89.8 Length: 2214
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 10 Gaps: 0

US-09-727-892a-99 (1-58) x AF326556 (1-2214)

Qy 16 LysGlyHisPheProHisGln 22
|||||
Db 281 AAGGCCATTTCACATCA 261

RESULT 35
AF326556/c 2235 bp mRNA linear ROD 02-DEC-2001
LOCUS Mus musculus DBA/2J phosphodiesterase 4B (Pde4b) mRNA, complete
DEFINITION cds.
ACCESSION AF326556
VERSION AF326556
KEYWORDS AF326556.1 GI:17225438
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2214)
REFERENCE Fehr, C., Belknap, J. K., Crabbe, J. C. and Buck, K. J.
AUTHORS High resolution mapping of a quantitative trait locus for acute
TITLE ethanol withdrawal on mouse chromosome 4 and characterization of
JOURNAL potential candidate genes
unpublished

REFERENCE 2 (bases 1 to 2235)
AUTHORS Fehr, C., Belknap, J. K., Crabbe, J. C. and Buck, K. J.
TITLE Direct Submission
JOURNAL Submitted (06-DEC-2000) Department of Behavioral Neuroscience,
Portland Alcohol Research Center, Oregon Health Sciences University
and Veterans Affairs Medical Center Portland (Research Services),
3710 SW US Veterans Hospital Road, Portland, OR 97201, USA

FEATURES
source
1..2235
/organism="Mus musculus"
/strain="DBA/2J"
/db_xref="taxon:10090"
/chromosome="4"
/map="46.8 cm"
/tissue_type="whole brain"
/dev_stage="adult"
1..2235
/gene="Pde4b"
23..2188
/note="cAMP specific"
/codon_start=1
/product="phosphodiesterase 4B"
/protein_id="AAL37402.1"
/db_xref="GI:17225439"
/translation="MTAKNSPKETASSESEVCITKEKQMRLELELPKLGQRPTSPK
ISPRSPRNSPCFFRKLLVNSITQRRTFYAHTCFVENGSPSGRSLPDPQAGSSSG
LVLHAAPPGHSQRESFLYRSDSDYDLSPKAMSNSSLPSQHGDDLIIVPPAVALS
LRVARNFTLLTNLHGAPNKRSPASQAPVRSVLSQESYOKLAMELELDLCLDL
ETIOTYRSVSEMANFKPKRMNRELTHSEMSRGNOVEYISNPLDKONDVIPS
TOKRREKKKOOLMTQISGVKLLHSSSLNNTSISRGVMTENEDHLAKLEEDLNKNG
LINFVAGYSHNRPLCTIMYAIROERDLTKFKISSPTPTMYMTLEDHHSVAYAN
SLRADVAOSTHYLSTPALDAVFTDLEILAIFPAALIHVDHGVSNQPLINTSEL
ALMINDSVLENHLLAVGFKLQEHCDIFQNLTKQKQRLRKWIDMVLATDMSKIM
SLADLKTWEETKKTSSGVLLDNYDRIOVLNMYCADLSNPTKSLLEYRQWTD
IMEEFFQGGKERERGMETSPMCDKHTASVEKSGVGFIDYIVHPLMETWADLVQPDQ
DIIDLTLEDNRNWQSMIPQSPPLDERSRDCGLMEKFOPELLEBEDSGPERKEGE
GHSYFSSTKLCVDPENRDSLEETDIDINTEKSPIDT"

BASE COUNT 650 a 560 c 530 g 495 t

ORIGIN

Alignment Scores:
Pred. No.: 90.5 Length: 2235
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 10 Gaps: 0

US-09-727-892a-99 (1-58) x AF326556 (1-2235)

Qy 16 LysGlyHisPheProHisGln 22
|||||
Db 281 AAGGCCATTTCACATCA 261

RESULT 36
AK025775 2233 bp mRNA linear PRI 29-SEP-2000
LOCUS Homo sapiens cDNA: FLJ22122 f1s, clone HEP19214.
DEFINITION AK025775
ACCESSION AK025775
VERSION AK025775.1 GI:10438393
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens hepatoma cell_line:Hepe2 cDNA to mRNA, clone_11b:HEP
clone:HEP19214.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (sites)
REFERENCE Kawabata, A., Haki, T., Kobatake, N., Inagaki, H., Ikema, Y.,
AUTHORS Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T.,
TITLE Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
JOURNAL NEBO human cDNA sequencing project
unpublished (2000)

REFERENCE 2 (bases 1 to 2333)
 AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
 Shibahara,T., Tanaka,T. and Nakamura,Y.
 TITLE Direct Submission
 JOURNAL Submitted (29-ANG-2000) to the DDBJ/EMBL/GenBank databases. Sumio
 Sugano, Institute of Medical Science, University of Tokyo,
 Laboratory of Genome Structure Analysis, Human Genome Center;
 Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
 (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
 Fax:81-3-5449-5416)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan; cDNA full insert
 sequencing: Research Association for Biotechnology; cDNA library
 construction, 5'- & 3'-end one pass sequencing: Department of
 Virology and Human Genome Center, Institute of Medical Science,
 University of Tokyo (partly supported by Science and Technology
 Agency).

FEATURES
 source
 1..2333
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_line="HepG2"
 /cell_type="hepatoma"
 /clone_id="HEP19214"
 /note="Cloning vector pME18SFL3"

BASE COUNT 551 a 488 c 585 g 709 t
 ORIGIN

Alignment Scores:
 Pred. No.: 93.7 Length: 2333
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 DB: Gaps: 0

US-09-727-892a-99 (1-58) x AK025775 (1-2333)

QY 33 Lysvalvaltyrserityrtyr 39
 |||||
 Db 1318 AAGTACTTTATCTTATAT 1338

RESULT 37
 AF202733 2433 bp mRNA linear ROD 15-NOV-2001
 LOCUS Rattus norvegicus CAMP-specific phosphodiesterase isoform PDE4B4
 DEFINITION (Pde4b) mRNA, complete cds.
 ACCESSION AF202733
 VERSION AF202733.1 GI:16930144
 KEYWORDS Norway rat.
 SOURCE Rattus norvegicus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 2433)
 Olsen,A.E. and Bolger,G.B.
 TITLE Novel PDE4B CAMP-specific phosphodiesterase isoforms and related
 technology
 JOURNAL Unpublished
 AUTHORS Olsen,A.E. and Bolger,G.B.
 TITLE Direct Submission
 JOURNAL Submitted (08-NOV-1999) Medicine (Oncology), University of Utah, 50
 North Medical Drive, Salt Lake City, UT 84132, USA

FEATURES
 source
 1..2433
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /tissue_type="cerebral cortex"
 1..2433
 gene

CDS
 /gene="Pde4b"
 262..2241
 /gene="Pde4b"
 /codon_start=1
 /product="CAMP-specific phosphodiesterase isoform PDE4B4"
 /protein_id="AA131764.1"
 /db_xref="GI:16930145"
 /translation="MLHVNDLPPRRHSWICEDVENGSPGSRSLDPPQASSSGGLVLA
 AAPGHSORESEFLYRSDDYDLSPRAMSRNSLSPSEQGDLLVYFPAQVLAISRY
 RNNFTLLTNHGAPNKRSPASQAPYRVLSQESYOKLAMELELDLQDLETIQ
 TYRSVSEMASNKRKRMNRELTHLSMSRSGNOVSERYISNTEFLDKONDEIPSPQKD
 REKKKQQLMTQISGVKTKMLHSSLSNNTSISRGVTEMDLAKELDELNKGWGNIF
 NVAGTSHNRPLTCTMTAIFQERDLATFEYSOTFVTYMTLEDHSDVAYHNSLHA
 ADVAOSTHVLSTPDLDAVFTDLIAAIFAAIHVDHPGVSNOFLINTNSALMT
 NDESVLENHMLAVGFKLQEBHCDIFQNTKRCQRTLRKRVIDVMTATDMSKMSILA
 DLKRWETKRVTSGLLDNYDRLOVRNMYCADLSNPKSLEYKQMTDRIMEE
 FPQGGKERERGRGEISPMCDKHTASVEKSOVGIDYIVHPIETNADLVQPPADQILD
 TLEDNRNRYDSMTPOSPSPFLDRSRDCGGLMEKRFELTLEBDSBEPKSGEGPNY
 FSSRTKYLCTVDIPENRDSLETDTIDIAETKSLIDT"

BASE COUNT 691 a 626 c 591 g 525 t
 ORIGIN

Alignment Scores:
 Pred. No.: 96.9 Length: 2433
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 DB: Gaps: 0

US-09-727-892a-99 (1-58) x AF202733 (1-2433)

QY 16 LysglYhispheprohISgln 22
 |||||
 Db 334 AAGGGCCATTTCCACATCAA 314

RESULT 38
 A06139
 LOCUS A06139
 DEFINITION Synthetic sequence of S fragment derived from capripoxvirus isolate
 RS-1.
 ACCESSION A06139
 VERSION A06139.1 GI:412755
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 REFERENCE 1 (bases 1 to 2497)
 AUTHORS
 TITLE VACCINE
 JOURNAL Patent: WO 9002190-A 1 08-MAR-1990;
 TITLE Location/Qualifiers
 source
 1..2497
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 1..153
 /gene="CF6"
 <1..153
 /gene="CF6"
 /codon_start=1
 /transl_table=1
 /protein_id="CAA00509.1"
 /db_xref="GI:412756"
 /translation="KIFISGLPNTWYSKALSINRQPIYKYCNDLLQSINSQGVSI
 NDIIRK"
 110..553
 /gene="CF7"
 110..553
 /gene="CF7"
 /codon_start=1
 /transl_table=1
 /protein_id="CAA00510.1"
 /db_xref="GI:412757"
 CDS
 gene

[illegible]

gene reveals a gene shared specifically with leporipoxvirus
J. Gen. Virol. 70 (pt 3), 525-533 (1989)
MEDLINE
COMMENT
The degrees of homology between the ORFs of KS-1 HindIII S and the
equivalent ORFs of other poxvirus genera are shown in [1].
FEATURES
source
1. .2497
/organism="Sheepox virus <strain KS-1>"
/strain="KS-1"
/isolate="KS-1"
/db_xref="taxon:10269"
/note="1 bp upstream of HindIII site.; isolated from Kenya
sheep"
<1. .153
/note="ORF CF6"
/codon_start=1
/protein_id="BAA00322.1"
/db_xref="GI:221121"
/translation="KLPIISGLPMTMSKSLSLNRPIITYKYCNLDLQSLNGSQVSI
NDILRK"
110. .553
/note="ORF CF7"
/codon_start=1
/protein_id="BAA00323.1"
/db_xref="GI:221122"
/translation="MDHSKYLITIFLENDSPFKYLSQDEDTAMSDIEITIVYTNFL
LSLIRSKDLKESIGYYEPLSECKTLVPSNNKNRILFNKIPIMILKQITVNG
YLSDFVTIMLKKELEFLESEPEPTIYIDPRDPFTILSLIHKH"
581. .1114
/note="ORF CF8"
/codon_start=1
/product="thymidine kinase"
/protein_id="BAA00324.1"
/db_xref="GI:221123"
/translation="MDGYIHLILIGPMESGKSTLIRKYOLAQYKCCVYAKLDI
RYGSAVYTHDNHVSALSTLILVYQKINPNFDIGDEQFPDIYSENNAMCK
IIITIALDSTFQREKFNPDILKLIFLSKVTIKLVANVCETKDAFAFRITKEIEILI
GGRKRYKSVCKCYFLE"
1186. .1779
/note="ORF CF8a"
/codon_start=1
/protein_id="BAA00325.1"
/db_xref="GI:221124"
/translation="WGLRHELDIILVSENLALKNVELLKDSIGCTINIKVNOOKID
FIILRPDMEVNRWKINWCVNGVVIDTLIKSFEEYSSSVVGFQVTFEFS
TSKKEEYPIVNTIKRYEIKDSMTCTINFESPISDQVYKDYINISDYLL
YDADDCIISDHDNDNMADDEEDDEVDIEDDYH"
1838. >2497
/note="ORF CF9"
/codon_start=1
/protein_id="BAA00326.1"
/db_xref="GI:221125"
/translation="MEAVSMDKPMYFDEIDNELEVDPKTSSEPKKLPYQGLKILL
CELPFSKLQHRGILDCCTIYVVSAGCTHIKYLDRFLMSGLVIRIILIDGRQHTI
LNGRLVDLTIRKFDSEYIRLKLKQLQSKTILISDVRSKRGNEPSTFLLSALQ
NIMVSLIKPAASSLKWRCPEPDQVWVXDFIIPHGNEMLOPPAPKYSAEIVNNIYSGNPI
KL"

RESULT 40
LOCUS SCYLRI42W 2554 bp DNA linear PLN 17-JUN-1997
DEFINITION S.cerevisiae chromosome XII reading frame ORF YLR142w.
ACCESSION Z73314
VERSION Z73314.1 GI:1360563
KEYWORDS
SOURCE baker's yeast.
ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetaceae; Saccharomycetaceae; Saccharomyces.
REFERENCE
1 (bases 1 to 2554)
AUTHORS Delius, H. and Hebling, U.
JOURNAL Unpublished
2 (bases 1 to 2554)
AUTHORS Rieger, M., Mueller-Auer, S. and Brueckner, M.
JOURNAL Unpublished
3 (bases 1 to 2554)
REFERENCE MIPS.
AUTHORS
JOURNAL
TITLE Direct Submission
SUBMITTED (22-MAY-1996) Data collected by MIPS on behalf of the
European yeast chromosome XII sequencing project. MIPS at the
Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a D-82152
Martinsried, FRG; E-mail: Mewes@mips.emblnet.org
FEATURES
source
1..2554
Location/Qualifiers
/organism="Saccharomyces cerevisiae"
/db_xref="taxon:4932"
/chromosome="XII"
412..1842
/gene="PURI"
412..1842
/gene="PURI"
/note="ORF YLR142w"
/codon_start=1
/protein_id="CAA97714.1"
/db_xref="GI:1360564"
/translaton="MIAKSSSLVTKRSRIPSLCPFLIKRSVSKTPHNSNTANLME
TPANAGNSVMAPNSINFLQTLPKKELFQGLGATLNSFLNTLTKLPYIRIP
VIEFFVSILYCGGNEKFEVIECGKRLQKRGSSNMALSTITNSBGTSLSTPDQIV
KETISSVHNLPIITGOLSKRIPNDIAIPGITALVNPHEVLNFSNPYKAO
RDOLIEKSSKIKTEIFELNOSLRLKYPKAPFVSTIDARKYDQENGVEYLGRIIF
OKNPSTSKLSCVGTQQLYLRDGDHILHKLQENGVLGKLVGAYINSEKNR
NOLIFGDKTGDENDRIITQVNDLINDSEYEGHLYVASHVYOSOMLVNTLKST
ODNYSKSNIVLIGLLGMADNVTVDLITNHGAKNIIKYPVGPPLFTKDYLLRLQEN
GDVRSNGPPLIAIKSIPKRYGL"
2358..2429
/gene="td(GTC)LR1 - systematic name"
/note="tRNA-Asp - common name; anticodon gene: GTC"
/product="tRNA-Asp"
2358..2429
/gene="td(GTC)LR1 - systematic name"
BASE COUNT 798 a 536 c 464 g 756 t
ORIGIN
Alignment Scores:
Pred. No.: 101 Length: 2554
Score: 7.00 Matches: 2554
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 8 Gaps: 0
US-09-727-892a-99 (1-58) x SCYLRI42W (1-2554)
QY 46 LysLysTyraLalYrLlelle 52
|||||
Db 1898 AAAAGATGATCATATATT 1918
RESULT 41
AF103948/c

LOCUS AF103948 2571 bp DNA linear PLN 01-APR-1999
DEFINITION Debaryomyces hanseni cytochrome P450 alkane hydroxylase (ALK1)
ACCESSION AF103948
VERSION AF103948.1 GI:4557161
KEYWORDS
SOURCE Debaryomyces hanseni.
ORGANISM Debaryomyces hanseni.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetaceae; Saccharomycetaceae; Debaryomyces.
REFERENCE
1 (bases 1 to 2571)
AUTHORS Yadav, J.S. and Loper, J.C.
JOURNAL Multiple P450alk (cytochrome P450 alkane hydroxylase) genes from
the halotolerant yeast Debaryomyces hanseni
Gene 226 (2), 139-146 (1999)
MEDLINE 9931473
PUBMED 9931473
REFERENCE 2 (bases 1 to 2571)
AUTHORS Yadav, J.S. and Loper, J.C.
JOURNAL Direct Submission
SUBMITTED (04-NOV-1998) Environmental Health, University of
Cincinnati, 231 Bethesda Ave., ML670056, Cincinnati, OH 45267-0056,
USA
FEATURES
source
1..2571
Location/Qualifiers
/organism="Debaryomyces hanseni"
/strain="ATCC 20317"
/db_xref="ATCC:20317"
/db_xref="taxon:4959"
700..2259
/gene="ALK1"
/note="CYP52A12"
700..2259
/gene="ALK1"
/note="member of CYP52 family"
/codon_start=1
/product="cytochrome P450 alkane hydroxylase"
/protein_id="A022536.1"
/db_xref="GI:4557162"
/translaton="MPDVDTVHNFTSWGIIAVALIGYHVPDYRIQIYMKKGV
PVRIDGPRPKLLYLSLKKKKGTLYNFKKEPFTVGDTSFRAGPVTSTNPE
NIRKLLATQSDRALGTRHAKQRPFLADGFTITDGGGKHSRMLRPOARQVAVK
SLEPHIOMLKKHARRAGAFDVQSLFRLTVDSATEFLGESVESLQDESIGKADA
VDEFGKAEAEVTAQVYLSISLQAKAFVNNKFFSSNEKVKRKFADYVYOKALN
SSPEELKEHQDQITFLYELVKQTRPQVHRLDQILILALAGRTVAGLSFFVEIAR
NPQWMLKEIEIEKEFGKGDARLEDTTFESLCKCYLALNLEVLALYSPQNEPV
AOKQTSIPRGSGPMDQPIRIARGQVTVYVYAMHBDQDFGKDSVEFPEPPEET
RKLGMAFLPENGGRICLGGQFALTEASYIAIALQDLPRLASHDEYPPKASHLTM
CHOSEVATISLA"
BASE COUNT 737 a 544 c 589 g 701 t
ORIGIN
Alignment Scores:
Pred. No.: 101 Length: 2571
Score: 7.00 Matches: 2571
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 8 Gaps: 0
US-09-727-892a-99 (1-58) x AF103948 (1-2571)
QY 43 LeuphehrlYsLysTyraLa 49
|||||
Db 1457 TTGTTACCAAAAGATGCT 1437
RESULT 42
AF1071067
LOCUS AY071067 2742 bp mRNA linear INV 20-DEC-2001
DEFINITION Dirosophila melanogaster REL14386 full length cDNA.
ACCESSION AY071067
VERSION AY071067.1 GI:17945263
KEYWORDS F1L_CDNA.

SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 2742)
AUTHORS Stapleton,M., Brokstein,P., Hong,L., Aghayani,A., Carlson,J.,
Champe,M., Chavez,C., Dorsett,V., Dresnek,D., Farfan,D., Frise,E.,
George,R., Gonzalez,M., Guarin,H., Krommiller,B., Li,P., Liao,G.,
Miranda,A., Mungall,C.J., Nunoo,J., Pacleb,J., Pachas,V., Park,S.,
Patel,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M.
and Celisner,S.
TITLE Direct Submission
JOURNAL Submitted (18-DEC-2001) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA
COMMENT Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila gene collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our Web site
(http://fruitfly.berkeley.edu) or send email to
cdna@fruitfly.berkeley.edu.
Location/Qualifiers
1..2742
/organism="Drosophila melanogaster"
/strain="y; cn by sp"
/db_xref="taxon:7227"
/map="91A2-91A2"
1..2742
/gene="CG7678"
/note="alignment with genomic scaffold AE003722"
/db_xref="FLYBASE:FBgn0038613"
86..2620
/gene="CG7678"
/note="Longest ORF"
/codon_start=1
/db_xref="FLYBASE:FBgn0038613"
/product="RE14386p"
/protein_id="AAL48689.1"
/db_xref="GI:17945264"
/translation="MSKMSGSGSNOESNIFRSEVMSLYOMYLOPEAAYDTIALGEV
GVOEFDLANKINAOKRFTGEVRKDELRIRYVTAELNKEGHVLDLDDPPAP
OPREIDLELHETKETEELIELANNVNLQTSYLESEMYQVLETPQFSDDSNHF
DLKMGTHRPEKSNGLHGFVAGTISRERAFRMLRISRGVFRKCDVDALTD
KLTGVNLKSVFVVEFGDQLQARIKRVGCFAMAMPCPSHSERQEMKVNTRLE
DLOVINQTSDRHCYLOALQKLPYSAWKKKGGIHTLNLNLEVDGSCILGSEW
VPRKRELEVALAAGSASVSPFINVLDTEKKEPTRETRKFPGRONLDLDAIG
IAGREVNPGIYCTITPPPLFAVMFGMGHTLFLGLMMVVIDEKLSKRGSEITNG
IFPAGRTIIMGLFAMYTGFHYNDISKSINVGRTWVYVNTYLTNPTLOLNS
VATRGYTPKGIIDPTMOSAKNITFLNTYKMKLSITFEVLAHRECVSVEHFYFKY
AYTILQFPOVLFLLMFYMCFFMYKWKVSTPTVEADTPGCASVLMFDWL
FKETALPGDNNVPEIQKNDLMLFVALLCIPIWILGRLYTKYORRNPAPRVE
VDEIVERKEIVTGGKEIITIEVAEHSQSGSEEDSEMSIHOAIHTLEYLSTIS
HTASYRLMALSLAHOLESVLTMTNVLAMGLONGYGAIGLEPIFVWPEFTAIWV
MMBSLSAFLTLRLHWEFMKSEYVGCGPTPSEFDILIVEDD"

BASE COUNT 666 a 669 c 738 g 669 t
ORIGIN

Alignment Scores:
Pred. No.: 107 Length: 2742

Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 3 Gaps: 0
US-09-727-892a-99 (1-58) x AY071067 (1-2742)
Qy 46 Lyslystyralatyrllelle 52
Db 1775 AAGAAATACGCTACATCATC 1795
RESULT 43
YSCPNT1 2881 bp DNA linear PLN 10-MAR-1994
LOCUS Yeast (S.cerevisiae) mitochondrial proline oxidase (PUT1) gene,
DEFINITION complete cds.
ACCESSION M18107
VERSION M18107.1 GI:172300
KEYWORDS proline oxidase.
SOURCE S.cerevisiae (strain 5288C) DNA, plasmid pMB8.
ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE 1 (bases 1 to 2881)
AUTHORS Wang,S.S. and Brandriss,M.C.
TITLE Proline utilization in Saccharomyces cerevisiae: sequence,
regulation, and mitochondrial localization of the PUT1 gene product
JOURNAL Mol. Cell. Biol. 7 (12), 4431-4440 (1987)
MEDLINE 88142835
COMMENT Draft entry and computer readable sequence for [1] kindly provided
by S.-S.Wang, 1-FEB-1988
FEATURES
source Location/Qualifiers
1..2881
/organism="Saccharomyces cerevisiae"
/db_xref="taxon:4932"
445..453
repeat_region
/note="direct repeat, copy A"
454..462
repeat_region
/note="direct repeat, copy B"
463..471
repeat_region
/note="direct repeat, copy C"
472..480
repeat_region
/note="direct repeat, copy D"
1012..2442
/note="proline oxidase"
/codon_start=1
/protein_id="AA16631.1"
/db_xref="GI:172301"
/translation="MMSKSLVTKRSIRSLCPPLIKRSYVSTPTPHSNRAMLME
TPRANFGNSVMAFPNSINFLQTLPRKELRDLGIGTATLNSFLNTIKLEPIIP
VIRFVSLYCGGNEKFEVLECGRLQKRGISNNMLSTLSENSGTSLSSTPDQIV
KETISSVHNILPNIIGOLSKPPTDAPGIALKPSALVDNPEHYVNFSPYKAO
RDOLIEKSKITKEIFELNOSLKLKYEERKAPFVSTIDAKRYDLOENGVEQLRIE
OKFNPSTSKISLCVGTMOIVLRDGGHILHLKLAONGYKLGKLVKRGVITSEKRN
NQLTFKGTGIDENVDITITQVVDILINEDSFGHLVYASHYQSGOMVTNLSKT
QDNSTAKSNIVLGLDGMADVITDITLTHGAKNITVTPWGPLETEKIDTLRLQEN
GDVAVSNGMPLAKIAKSIKPRGL"
BASE COUNT 890 a 647 c 571 g 773 t
ORIGIN 2 bp upstream of ClaI site; Chromosome 12.
Alignment Scores:
Pred. No.: 111 Length: 2881
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 8 Gaps: 0
US-09-727-892a-99 (1-58) x YSCPNT1 (1-2881)
Qy 46 Lyslystyralatyrllelle 52
|||||

Db 2498 AAAAGTATGATCATTAATT 2518

RESULT 44

AF208023/C

LOCUS AF208023 2912 bp mRNA linear ROD 20-MAR-2001
DEFINITION Mus musculus cAMP-specific phosphodiesterase (Pde4b) mRNA, complete cds.

ACCESSION AF208023 GI:8901296

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

Db 476 AAGGCCATTTCCACATCAA 456
RESULT 45
AF297397/C
LOCUS AF297397 2912 bp mRNA linear ROD 07-JUL-2000
DEFINITION Mus musculus mRNA for phosphodiesterase 4B, cAMP specific, isoform 3 (Pde4b gene).
ACCESSION AF297397 GI:8979836
VERSION AJ297397.1
KEYWORDS alternative splicing; Pde4b gene; phosphodiesterase 4B, cAMP specific.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 2912)
AUTHORS Cherry, J.A., Thompson, B.E. and Pho, V.
TITLE Diazepam and rolipram differentially inhibit cAMP-specific phosphodiesterases PDE4A1 and PDE4B3 in the mouse
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2912)
AUTHORS Cherry, J.A.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-2000) Cherry J.A., psychology, Boston University, 64 Cummington Street, Boston, MA 02215, USA
location/Qualifiers
1. .2912
/organism="Mus musculus"
/strain="Swiss Webster"
/db_xref="taxon:10090"
/country="USA"
218. .2383
/gene="Pde4b"
218. .2383
/note="isoform 3"
/evidence="experimental"
/product="phosphodiesterase 4B, cAMP specific"
/protein_id="CA966770.1"
/db_xref="GI:8979837"
/translation="MTAKNSPKFEFTASESEVCIKTFKQMRLELPLKPLGNPTSPK
ISPRSSPNSPCFRRKLIVNKSIRORRRPTVAHCTFDVNGSPSPRSLDPDQAGSSG
LVYHAAPGHSQRESFLYDSDYDLSFKAMRSNLSPEQHGDDLIYTFPAVILAS
LTVRNNFTLNLHGAPNKRSPASQAPSVRSLOEESYOKLAMELELDLCLDOL
ETQDRKKKKQOOLMTQISYVKKLHSSSLNNTSISRFQINTENEDHLAKLEDKAKK
LNFENAGVSHNPLTCIMYATFOERDLTKRISSDTFVYMTLEDHYSVAVYHN
SLAAVAVQSTHYHLSTPALADVFTDLLETLAIFAAIHVDHDPGVSNQPLINTSEL
ALMYNDESVLENHHLAVGFKLQEBHCDIFONTKROTLRKMYIDMYLATMYSKHM
SLADLKTIVETKKTSSGVLNDNTDRIOVRNWHCHDLSNPKSLELYRQWTDK
IMEFFQDGKREKRGEMELSPMCKHTASVEKQVGFIDYIVHPLMETWADLVQPDQO
DILDTLEDNRNYSMIPOSPSPIDERSDCGLMEKQFELTLEEDSGPEKEGE
GHYSFSTTLTCYIDPNDRLSEETDIDATBDKSPIDT"BASE COUNT 841 a 721 c 681 g 669 t
ORIGIN
Alignment Scores:
Pred. No.: 112 Length: 2912
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
Gaps: 0
US-09-727-892a-99 (1-58) x MMU297397 (1-2912)
Cy 16 LysGlyHisPheProHisGln 22
Db 476 AAGGCCATTTCCACATCAA 456BASE COUNT 841 a 721 c 681 g 669 t
ORIGIN
Alignment Scores:
Pred. No.: 112 Length: 2912
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
Gaps: 0
US-09-727-892a-99 (1-58) x AF208023 (1-2912)
Cy 16 LysGlyHisPheProHisGln 22
Db 476 AAGGCCATTTCCACATCAA 456

Cy 16 LysGlyHisPheProHisGln 22

Wed Nov 6 12:59:55 2002

us-09-727-892a-99.NAolig.rge

Page 33

Search completed: November 5, 2002, 05:27:40
Job time : 2074 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 04:25:09 ; Search time 62 Seconds

(without alignments)
103.908 Million cell updates/sec

Title: US-09-727-892a-99

Perfect score: 58
Sequence: 1 MERKRYTVLLYCEIKGHP.....YEVNLFTRKAYIIEIKEL 58

Scoring table: OLIGO
Gapop 60.0 , Gapept 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

A.Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	100.0	58	21	AA16557
2	7	12.1	677	22	AB161419
3	7	12.1	844	22	AB171568
4	7	12.1	2424	22	AB168256
5	6	10.3	15	20	AA107920
6	6	10.3	45	22	AA174575
7	6	10.3	61	21	AA103465
8	6	10.3	79	20	AA159332
9	6	10.3	109	22	AA1030542
10	6	10.3	110	21	AA164580
11	6	10.3	120	22	AA1604212

12	6	10.3	132	21	AA105492	Arabidopsis thalia
13	6	10.3	169	21	AA136625	Arabidopsis thalia
14	6	10.3	178	22	AA105778	Soybean invertase
15	6	10.3	180	22	AA125554	Human G Protein-Co
16	6	10.3	186	22	AA166487	putative P. abyssal
17	6	10.3	188	22	AA117158	Novel signal trans
18	6	10.3	195	21	AA133164	Eucalyptus grandis
19	6	10.3	196	21	AA108657	Arabidopsis thalia
20	6	10.3	223	22	AB117135	Human nervous syst
21	6	10.3	233	22	AA119799	Human novel extrac
22	6	10.3	235	21	AA132552	Eucalyptus grandis
23	6	10.3	240	22	AA118242	Plasmodium falcipa
24	6	10.3	240	22	AA193486	Human protein sequ
25	6	10.3	241	21	AA139149	Human secreted pro
26	6	10.3	247	21	AA105491	Arabidopsis thalia
27	6	10.3	248	21	AA105490	Arabidopsis thalia
28	6	10.3	251	22	AA164578	Drosophila melanog
29	6	10.3	262	21	AA145775	Arabidopsis thalia
30	6	10.3	262	22	AA158183	Drosophila melanog
31	6	10.3	262	22	AA163475	Drosophila melanog
32	6	10.3	268	22	AA167569	Amino acid sequenc
33	6	10.3	289	22	AA130285	Novel human diagno
34	6	10.3	293	21	AA153344	Human colon cancer
35	6	10.3	294	21	AA141745	Arabidopsis thalia
36	6	10.3	296	21	AA141744	Arabidopsis thalia
37	6	10.3	298	20	AA173880	Human prostate tum
38	6	10.3	298	21	AA143779	Human cancer assoc
39	6	10.3	298	22	AA123487	Novel human enzyme
40	6	10.3	298	22	AA125792	Human protein sequ
41	6	10.3	322	21	AA108656	Arabidopsis thalia
42	6	10.3	336	22	AA125613	Human G Protein-Co
43	6	10.3	336	22	AA108556	Human seven trans
44	6	10.3	336	22	AA104374	Human G-protein co
45	6	10.3	336	22	AA164299	Human GTP-binding

ALIGNMENTS

RESULT 1	AA16557	standard; Protein; 58 AA.
ID	AA16557;	
AC	AA16557;	
XX	27-OCT-2000	(first entry)
XX		
DE	Bacteriophage 44AHD protein sequence 44AHDORF025.	
XX		
KW	Bacteriophage; antimicrobial; genome; identification; antibacterial;	
KW	bacterial growth inhibition; bacterial infection.	
XX		
OS	Bacteriophage 44AHD.	
XX		
PN	W0200032825-A2.	
XX		
PD	08-JUN-2000.	
XX		
PF	03-DEC-1999;	99WO-IB02040.
XX		
PR	03-DEC-1998;	98US-0110992.
PR	03-JUN-1999;	99US-0326144.
PR	28-SEP-1999;	99US-0407804.
PR	30-SEP-1999;	99US-0157218.
PR	01-DEC-1999;	99US-0168777.
PR	02-DEC-1999;	99US-0454252.
XX		
PA	(PHAG-) PHAGETECH INC.	
XX		
PI	Pelletier J, Gros P, Dubow M;	
XX		
DR	WPI: 2000-412361/35.	
	N-PSDB; AAA69042.	

XX Identifying a bacteriophage coding region for treating bacterial
PT infections comprises identifying a nucleic acid encoding a product that
PT inhibits bacteria when a bacteriophage infects a bacterium
XX
PS Example 9; Page 278; 456pp; English.
XX
CC The present invention describes a method for identifying a bacteriophage
CC coding region encoding a product active on an essential bacterial
CC target. The method comprises identifying a nucleic acid sequence encoding
CC a gene product that provides a bacteria-inhibiting function when an
CC uncharacterised bacteriophage infects a pathogenic bacterium. The
CC compound active on a target of a bacteriophage inhibitor protein in a
CC bacteria is used to treat or prevent a bacterial infection in an animal.
CC AAA68243 to AAA69442 and AAB16523 to AAB16954 represent bacteriophage
CC nucleotide and protein sequences which are used in the exemplification of
CC the present invention.
XX
SQ Sequence 58 AA;
XX
Query Match 100.0%; Score 58; DB 21; Length 58;
Best Local Similarity 100.0%; Pred. No. 4.2e-54;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 MERRKYTVLYCDEIKGFPHQISMFDLYDAKVVSYYEXNLTFTKRYATIEYIKET 58
1 MERRKYTVLYCDEIKGFPHQISMFDLYDAKVVSYYEXNLTFTKRYATIEYIKET 58
Db 1 MERRKYTVLYCDEIKGFPHQISMFDLYDAKVVSYYEXNLTFTKRYATIEYIKET 58
XX
RESULT 2
AAB61419
ID ABB61419 standard; Protein: 677 AA.
XX
AC ABB61419;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 11049.
XX
KM Drosophila: developmental biology; cell signalling; insecticide;
KM pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PE 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI: 2001-656860/75.
DR N-PSDB; ABL05522.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 11049; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins

CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 677 AA;
XX
Query Match 12.1%; Score 7; DB 22; Length 677;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 10 LYCDEIK 16
Db 276 LYCDEIK 282
XX
RESULT 3
AAB71568
ID ABB71568 standard; Protein: 844 AA.
XX
AC ABB71568;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 41496.
XX
KM Drosophila: developmental biology; cell signalling; insecticide;
KM pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PE 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI: 2001-656860/75.
DR N-PSDB; ABL15671.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 41496; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 844 AA;
XX
Query Match 12.1%; Score 7; DB 22; Length 844;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 46 KRYAYIT 52
1111111111

Db 564 KRYAVII 570

RESULT 4

ABB68256

ID ABB68256 standard; Protein; 2424 AA.

XX ABB68256;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 31560.

XX Drosophila: developmental biology; cell signalling; insecticide;

XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI: 2001-656860/75.

XX N-PSDB; ABL12359.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions -

XX Disclosure; SEQ ID NO 31560; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

XX sequences (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2424 AA;

XX

XX Query Match 12.1%; Score 7; DB 22; Length 2424;

XX Best Local Similarity 100.0%; Pred. No. 1.1e+02;

XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

XX 27 EDLYDAK 33

XX 1206 EDLYDAK 1212

XX

XX RESULT 5

XX AAY07920

XX ID AAY07920 standard; Protein; 15 AA.

XX AAY07920;

XX 06-JUL-1999 (first entry)

XX Human secreted protein fragment encoded from gene 69.

XX Human; secreted protein; treatment; prevention; protein therapy; AIDS;

XX gene therapy; diagnosis; cancer; tumour; neurodegenerative disorder;

XX developmental abnormality; fetal deficiency; blood disorder; leukemia;

XX immune system disease; autoimmune disease; hepatic disease; lymphoma;

XX renal disease; inflammation; allergy; Alzheimer's disease; schizophrenia;

XX cognitive disorder; prostate disease; skeletal; cardiac; muscle disorder;

XX pulmonary disorder; transplant rejection; osteoporosis;

XX arthritis; malignancy; digestive; endocrine; infection.

XX Homo sapiens.

XX WO9918208-A1.

XX 15-APR-1999.

XX 01-OCT-1998; 98WO-US20775.

XX 02-OCT-1997; 97US-0060884.

XX 02-OCT-1997; 97US-0060833.

XX 02-OCT-1997; 97US-0060836.

XX 02-OCT-1997; 97US-0060837.

XX 02-OCT-1997; 97US-0060838.

XX 02-OCT-1997; 97US-0060839.

XX 02-OCT-1997; 97US-0060843.

XX 02-OCT-1997; 97US-0060862.

XX 02-OCT-1997; 97US-0060866.

XX 02-OCT-1997; 97US-0060874.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Carter KC, Duan DR, Endress GA, Feng P, Ferrie AM;

XX Florence KA, Greene JM, Janat F, Lafleur DW, Ni J;

XX Rosen CA, Ruben SM, Shi Y, Young P, Yu G;

XX WPI: 1999-264022/22.

XX N-PSDB; AAX37519.

XX New isolated human genes and the secreted polypeptides they encode

XX Claim 1b; Page 310; 368pp; English.

XX This invention describes novel isolated human genes and the secreted

XX proteins they encode. The products of the invention are useful for

XX preventing, treating or ameliorating medical conditions, e.g. by protein

XX or gene therapy. Also pathological conditions can be diagnosed by

XX determining the amount of the new polypeptides in a sample or by

XX determining the presence of mutations in the new polynucleotides.

XX Specific uses are described for each of the 101 polynucleotides, based

XX on which tissues they are most highly expressed in, and include

XX developing products for the diagnosis or treatment of cancer, tumours,

XX neurodegenerative disorders, developmental abnormalities and fetal

XX deficiencies, blood disorders, leukemias, diseases of the immune system,

XX autoimmune diseases, hepatic and renal disease, lymphomas, inflammation,

XX allergies, Alzheimer's and cognitive disorders, schizophrenia, prostate

XX disease, skeletal or cardiac muscle disorders, pulmonary disorders,

XX transplant rejection, disorders involving osteoclasts such as

XX osteoporosis, arthritis or malignancies, digestive/endocrine disorders,

XX infections and AIDS. The human secreted proteins of the invention are

XX represented in AAX37451-X37552.

XX

XX Sequence 15 AA;

XX

XX Query Match 10.3%; Score 6; DB 20; Length 15;

XX Best Local Similarity 100.0%; Pred. No. 14;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

XX 5 YKTVLL 10

XX 4 YKTVLL 9

XX

XX RESULT 6

AA674575
ID AAG74575 standard; Protein; 45 AA.
XX
XX
AC AAG74575;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:5339.
XX
KW Human: colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma.
XX
OS Homo sapiens.
XX
PN MO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI: 2001-235357/24.
DR N-PSDB; AAH33980.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS Claim 11; Page 6988; 9803pp; English.
XX
CC AAH32943 to AAH7195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 45 AA:

Query Match 10.3%; Score 6; DB 22; Length 45;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 HQISMF 26
 |||||
Db 16 HQISMF 21

RESULT 7
AAG03465
ID AAG03465 standard; Protein; 61 AA.
XX
AC AAG03465;
XX
DT 06-OCT-2000 (first entry)

XX
DE Human secreted protein, SEQ ID NO: 7546.
XX
XX
KW Human: 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
PI Dunas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI: 2000-500381/45.
DR N-PSDB; AAC03471.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX
PS Claim 13; SEQ ID 7546; 71pp + CD-ROM; English.
XX
CC The present sequence is a polypeptide encoded by one of a large number
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC were prepared from total human RNAs or polyA+ RNAs derived from 30
CC different tissues. EST sequences usually correspond mainly to the 3'
CC untranslated region (UTR) of the mRNA because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5'
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC ends and can therefore be used to obtain full length cDNAs and genomic
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.
XX
SQ Sequence 61 AA:

Query Match 10.3%; Score 6; DB 21; Length 61;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 YNLFTR 46
 |||||
Db 8 YNLFTR 13

RESULT 8
AAY59932
ID AAY59932 standard; Protein; 79 AA.
XX
AC AAY59932;
XX
DT 28-JAN-2000 (first entry)
XX
DE Human myometrium tumour EST encoded protein 12.
XX
KW Myometrium; tumour; human; expressed sequence tag; EST; uterine myoma;
KW treatment; carcinoma; cancer; gene therapy.
XX
OS Homo sapiens.
XX
PN DE1981947-A1.
XX
DT 28-OCT-1999.

PF 17-APR-1998; 98DE-1017947.
 XX
 PR 17-APR-1998; 98DE-1017947.
 XX
 PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX
 PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Palarsky C, Dahl E;
 XX WPI: 1999-602380/52.
 DR N-PSDB; AA241967.
 XX
 PT New nucleic acid sequences expressed in uterine myoma, and derived
 PT polypeptides, for treatment of uterine carcinoma and identification of
 PT therapeutic agents -
 XX
 PS Claim 23; Page 70; 86pp; German.
 XX
 CC This invention describes novel polypeptide sequences (I), fragments of
 CC (I) fragments and their encoding nucleic acids (II) which are highly
 CC expressed in human uterine myoma. (II) are used for recombinant
 CC expression of (I) and to isolate complete genes. (I) are used to
 CC identify agents suitable for treatment of uterine carcinoma, to directly
 CC treat this form of cancer (including expression from gene therapy
 CC vectors) and are used in a preparation for cancer treatment. (I) is also
 CC used for the generation of specific antibodies. (II) are identified by
 CC assembling ESTs (expressed sequence tags) from a particular tissue type
 CC before comparison of expression patterns. This allows a significantly
 CC longer fragment of the gene to be revealed and therefore reduces the
 CC number of failures associated with the fact that ESTs from different
 CC libraries may represent different parts of the same unknown gene,
 CC distorting the estimated frequency of occurrence in a particular tissue.
 CC AAY59921-Y59940 represent protein fragments encoded by the human
 CC myometrium tumour CDNA library derived EST fragments represented in
 CC AA241950-Z41980.
 CC
 XX
 SQ Sequence 79 AA;
 XX
 Query Match 10.3%; Score 6; DB 20; Length 79;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 TVLLYC 12
 II
 DB 55 TVLLYC 60
 XX
 RESULT 9
 AAU30542
 ID AAU30542 standard; Protein; 109 AA.
 XX
 AC AAU30542;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secreted protein #1033.
 XX
 KW Human: vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX
 OS Homo sapiens.
 OS
 PN WO200179449-A2.
 PD 25-OCT-2001.
 XX
 PF 16-APR-2001; 2001WO-US08656.
 XX
 PR 18-APR-2000; 2000US-0552929.
 PR 26-JAN-2001; 2001US-0770160.
 XX
 PA (HYSE-) HYSEQ INC.
 XX

PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI: 2001-611725/70.
 XX
 PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy -
 XX
 PS Claim 20; Page 307; 765pp; English.
 XX
 CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.
 CC
 XX
 SQ Sequence 109 AA;
 XX
 Query Match 10.3%; Score 6; DB 22; Length 109;
 Best Local Similarity 100.0%; Pred. No. 80;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 VLYXCD 13
 IIIII
 DB 30 VLYXCD 35
 XX
 RESULT 10
 AAY64580
 ID AAY64580 standard; Peptide; 110 AA.
 XX
 AC AAY64580;
 XX
 DT 02-MAR-2000 (first entry)
 XX
 DE Nonclassical cadherin extracellular domain SEQ ID NO:8.
 XX
 KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
 KW neurological disease.
 XX
 OS Mammalia.
 OS
 PN WO9957149-A2.
 PD 11-NOV-1999.
 XX
 PF 05-MAY-1999; 99WO-CA00363.
 XX
 PR 05-MAY-1998; 98US-0073040.
 PR 06-NOV-1998; 98US-0187859.
 PR 20-JAN-1999; 99US-0234395.
 PR 08-MAR-1999; 99US-0264516.
 XX
 PA (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX
 PI Blaschuk OW, Gour BJ, Byers S;

XX WPI: 2000-038791/03.
DR New cadherin modulating agents, used for modulating nonclassical
XX cadherin-mediated functions for treating e.g. cancers, obesity,
PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological
PT disease
XX
PS Disclosure: Fig 2; 252pp; English.
XX
XX The present invention describes cadherin modulating agents (MA)
CC comprising peptides which comprise a nonclassical cadherin cell adhesion
CC recognition (CAR) sequence. The MAs can be used for modulating
CC nonclassical cadherin-mediated functions. They can be used for e.g.
CC inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC mammal, enhancing delivery of a drug through the skin of a mammal,
CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC expressing cell, preventing or treating obesity in a mammal, stimulating
CC blood vessel regression in a mammal, enhancing drug delivery to the
CC central nervous system, treating a demyelinating neurological disease,
CC increasing vasopermeability in a mammal, enhancing adhesion of
CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in
CC a mammal, or preventing pregnancy in a mammal. They can also be used for
CC e.g. enhancing or directing neurite outgrowth, facilitating wound
CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue
CC in a mammal. They can also be used for treating e.g. psoriasis,
CC arthritis, age-related macular degeneration, multiple sclerosis and
CC diabetes. The products can also be used for detection and diagnosis and
CC in bioreactors. AAY60592 to AAY64572 represent specifically claimed
CC peptides, and AAY64573 to AAY64643 and AA233183 to AA233186 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 110 AA:

Query Match 10.3%; Score 6; DB 21; Length 110;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 36 AKVYVS 41
QY 32 AKVYVS 37
111111
36 AKVYVS 41
RESULT 11
AAE04212
ID AAE04212 standard; Protein: 120 AA.
AC AAE04212;
XX
DT 09-ANG-2001 (first entry)
XX
DE Human gene 15 encoded secreted protein HMBN029, SEQ ID NO:67.
XX
XX Human: secreted protein; proliferative disorder; cancer; tumour;
KM foetal abnormality; developmental abnormality; hemotopoietic disorder;
KM immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KM inflammation; allergy; neurological disorder; Alzheimer's disease;
KM Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KM skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KM cardiovascular disorder; angiogenic disorder; kidney disorder;
KM gastrointestinal disorder; pregnancy-related disorder; gene therapy;
KM endocrine disorder; infection; wound healing; vulnery;
KM cell culture; chemotaxis; food additive; chromosome 14;
KM binding partner identification.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Peptide 1..1
FT /label= signal_peptide
FT Protein 2..120

FT /note= "Mature secreted protein"
EN W0200136432-A2.
XX
XX 25-MAY-2001.
ED
XX
XX 15-NOV-2000; 2000WO-0531162.
XX
XX 19-NOV-1999; 99US-0166415.
XX 30-JUN-2000; 2000US-0215136.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Komatsoulis GA, Baker KP, Young PE;
DR WPI: 2001-343793/36.
DR N-PSDB; AAD08502.
XX
XX Isolated nucleic acid molecule encoding a human secreted protein is
XX used in preventing, treating or ameliorating a medical condition
XX
XX Claim 11; Page 445; 509pp; English.
XX
XX AAD08488-AAD08529 represent cDNAs corresponding to 18 human secreted
XX protein genes, and AAE04199-AAE04239 represent the proteins they encode.
XX AAE04240-AAE04297 represent human secreted protein fragments or variants.
XX The secreted proteins and their genes are useful for preventing,
XX treating or ameliorating medical conditions, e.g., by protein or gene
XX therapy. Pathological conditions can be diagnosed by determining the
XX amount of the new protein in a sample or by determining the presence of
XX mutations in the new genes. Specific uses are described for each of the
XX 18 genes, based on the tissues in which they are most highly expressed,
XX and include developing products for the diagnosis or treatment of
XX proliferative disorders, cancer, tumours, foetal and developmental
XX abnormalities, haematopoietic disorders, diseases of the immune system,
XX AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
XX allergies, neurological disorders (e.g., Alzheimer's disease,
XX Parkinson's disease), cognitive disorders, schizophrenia, asthma,
XX skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
XX cardiovascular disorders, angiogenic disorders, kidney disorders,
XX gastrointestinal disorders, pregnancy-related disorders, endocrine
XX disorders, and infections. The proteins can also be used to aid wound
XX healing and epithelial cell proliferation, to prevent skin aging due to
XX sunburn, to maintain organs before transplantation, for supporting cell
XX culture of primary tissues, to regenerate tissues, to identify their
XX cognate ligands or binding partners, and in chemotaxis, and can be used
XX as a food additive or preservative to modify storage properties.
XX Antibodies specific for a protein of the invention can be used in
XX alleviating symptoms associated with the disorders mentioned above, and
XX in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
XX immunosorbent assay (ELISA). The present sequence represents a human
XX secreted protein of the invention.
XX
SQ Sequence 120 AA:

Query Match 10.3%; Score 6; DB 22; Length 120;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 17 GHFPHQ 22
111111
11 GHFPHQ 16
RESULT 12
AAG05492
ID AAG05492 standard; Protein: 132 AA.
AC AAG05492;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 1917.

XX	Protein identification: signal transduction pathway; metabolic pathway;	PR	23-JUN-1999;	99US-0140353.
KW	hybridisation assay: genetic mapping; gene expression control; promoter;	PR	23-JUN-1999;	99US-0140354.
KW	termination sequence.	PR	24-JUN-1999;	99US-0140685.
XX		PR	28-JUN-1999;	99US-0140823.
OS	Arabidopsis thaliana.	PR	29-JUN-1999;	99US-0140991.
XX		PR	30-JUN-1999;	99US-0141287.
PN	EPI033405-A2.	PR	01-JUL-1999;	99US-0141842.
XX		PR	01-JUL-1999;	99US-0142154.
PD		PR	02-JUL-1999;	99US-0142055.
XX		PR	06-JUL-1999;	99US-0142380.
XX		PR	08-JUL-1999;	99US-0142803.
PF	06-SEP-2000.	PR	09-JUL-1999;	99US-0142920.
XX		PR	12-JUL-1999;	99US-0142977.
XX		PR	13-JUL-1999;	99US-0143542.
PR	25-FEB-2000; 2000EP-0301439.	PR	14-JUL-1999;	99US-0143624.
PR		PR	15-JUL-1999;	99US-0144005.
PR		PR	16-JUL-1999;	99US-0144085.
PR		PR	19-JUL-1999;	99US-0144086.
PR		PR	19-JUL-1999;	99US-0144331.
PR		PR	19-JUL-1999;	99US-0144332.
PR		PR	19-JUL-1999;	99US-0144333.
PR		PR	19-JUL-1999;	99US-0144334.
PR		PR	20-JUL-1999;	99US-0144352.
PR		PR	20-JUL-1999;	99US-0144632.
PR		PR	20-JUL-1999;	99US-0144884.
PR		PR	21-JUL-1999;	99US-0144814.
PR		PR	21-JUL-1999;	99US-0145086.
PR		PR	21-JUL-1999;	99US-0145088.
PR		PR	22-JUL-1999;	99US-0145085.
PR		PR	22-JUL-1999;	99US-0145087.
PR		PR	22-JUL-1999;	99US-0145089.
PR		PR	22-JUL-1999;	99US-0145192.
PR		PR	23-JUL-1999;	99US-0145145.
PR		PR	23-JUL-1999;	99US-0145224.
PR		PR	26-JUL-1999;	99US-0145276.
PR		PR	27-JUL-1999;	99US-0145913.
PR		PR	27-JUL-1999;	99US-0145918.
PR		PR	28-JUL-1999;	99US-0145919.
PR		PR	02-AUG-1999;	99US-0146386.
PR		PR	02-AUG-1999;	99US-0146388.
PR		PR	03-AUG-1999;	99US-0146389.
PR		PR	04-AUG-1999;	99US-0147038.
PR		PR	04-AUG-1999;	99US-0147204.
PR		PR	05-AUG-1999;	99US-0147302.
PR		PR	05-AUG-1999;	99US-0147302.
PR		PR	06-AUG-1999;	99US-0147260.
PR		PR	06-AUG-1999;	99US-0147303.
PR		PR	06-AUG-1999;	99US-0147416.
PR		PR	09-AUG-1999;	99US-0147493.
PR		PR	10-AUG-1999;	99US-0148171.
PR		PR	11-AUG-1999;	99US-0148319.
PR		PR	12-AUG-1999;	99US-0148341.
PR		PR	13-AUG-1999;	99US-0148565.
PR		PR	16-AUG-1999;	99US-0149368.
PR		PR	17-AUG-1999;	99US-0149175.
PR		PR	18-AUG-1999;	99US-0149426.
PR		PR	20-AUG-1999;	99US-0149722.
PR		PR	20-AUG-1999;	99US-0149723.
PR		PR	23-AUG-1999;	99US-0149929.
PR		PR	23-AUG-1999;	99US-0149930.
PR		PR	25-AUG-1999;	99US-0150566.
PR		PR	26-AUG-1999;	99US-0150884.
PR		PR	27-AUG-1999;	99US-0151065.
PR		PR	27-AUG-1999;	99US-0151066.
PR		PR	30-AUG-1999;	99US-0151080.
PR		PR		99US-0151303.

XX		PR	23-JUN-1999;	99US-0140353.
KW	hybridisation assay: genetic mapping; gene expression control; promoter;	PR	23-JUN-1999;	99US-0140354.
KW	termination sequence.	PR	24-JUN-1999;	99US-0140685.
XX		PR	28-JUN-1999;	99US-0140823.
OS	Arabidopsis thaliana.	PR	29-JUN-1999;	99US-0140991.
XX		PR	30-JUN-1999;	99US-0141287.
PN	EPI033405-A2.	PR	01-JUL-1999;	99US-0141842.
XX		PR	01-JUL-1999;	99US-0142154.
PD		PR	02-JUL-1999;	99US-0142055.
XX		PR	06-JUL-1999;	99US-0142380.
XX		PR	08-JUL-1999;	99US-0142803.
PF	06-SEP-2000.	PR	09-JUL-1999;	99US-0142920.
XX		PR	12-JUL-1999;	99US-0142977.
XX		PR	13-JUL-1999;	99US-0143542.
PR	25-FEB-2000; 2000EP-0301439.	PR	14-JUL-1999;	99US-0143624.
PR		PR	15-JUL-1999;	99US-0144005.
PR		PR	16-JUL-1999;	99US-0144085.
PR		PR	19-JUL-1999;	99US-0144086.
PR		PR	19-JUL-1999;	99US-0144331.
PR		PR	19-JUL-1999;	99US-0144332.
PR		PR	19-JUL-1999;	99US-0144333.
PR		PR	19-JUL-1999;	99US-0144334.
PR		PR	20-JUL-1999;	99US-0144352.
PR		PR	20-JUL-1999;	99US-0144632.
PR		PR	20-JUL-1999;	99US-0144884.
PR		PR	21-JUL-1999;	99US-0144814.
PR		PR	21-JUL-1999;	99US-0145086.
PR		PR	21-JUL-1999;	99US-0145088.
PR		PR	22-JUL-1999;	99US-0145085.
PR		PR	22-JUL-1999;	99US-0145087.
PR		PR	22-JUL-1999;	99US-0145089.
PR		PR	22-JUL-1999;	99US-0145192.
PR		PR	23-JUL-1999;	99US-0145145.
PR		PR	23-JUL-1999;	99US-0145224.
PR		PR	26-JUL-1999;	99US-0145276.
PR		PR	27-JUL-1999;	99US-0145913.
PR		PR	27-JUL-1999;	99US-0145918.
PR		PR	28-JUL-1999;	99US-0145919.
PR		PR	02-AUG-1999;	99US-0146386.
PR		PR	02-AUG-1999;	99US-0146388.
PR		PR	03-AUG-1999;	99US-0146389.
PR		PR	04-AUG-1999;	99US-0147038.
PR		PR	04-AUG-1999;	99US-0147204.
PR		PR	05-AUG-1999;	99US-0147302.
PR		PR	05-AUG-1999;	99US-0147302.
PR		PR	06-AUG-1999;	99US-0147260.
PR		PR	06-AUG-1999;	99US-0147303.
PR		PR	06-AUG-1999;	99US-0147416.
PR		PR	09-AUG-1999;	99US-0147493.
PR		PR	10-AUG-1999;	99US-0148171.
PR		PR	11-AUG-1999;	99US-0148319.
PR		PR	12-AUG-1999;	99US-0148341.
PR		PR	13-AUG-1999;	99US-0148565.
PR		PR	16-AUG-1999;	99US-0149368.
PR		PR	17-AUG-1999;	99US-0149175.
PR		PR	18-AUG-1999;	99US-0149426.
PR		PR	20-AUG-1999;	99US-0149722.
PR		PR	20-AUG-1999;	99US-0149723.
PR		PR	23-AUG-1999;	99US-0149929.
PR		PR	23-AUG-1999;	99US-0149930.
PR		PR	25-AUG-1999;	99US-0150566.
PR		PR	26-AUG-1999;	99US-0150884.
PR		PR	27-AUG-1999;	99US-0151065.
PR		PR	27-AUG-1999;	99US-0151066.
PR		PR	30-AUG-1999;	99US-0151080.
PR		PR		99US-0151303.

PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161922.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

Query Match 10.3%; Score 6; DB 21; Length 132;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 EYKEI 58
DB 42 EYKEI 47

RESULT 13
AAAG33625
ID AAAG33625 standard; Protein; 169 AA.

AC AAAG33625;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 40776.

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.
XX 06-SEP-2000.
PD 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 9905-0121825.
PR 05-MAR-1999; 9905-0123180.
PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 08-APR-1999; 9905-0128714.
PR 16-APR-1999; 9905-0129845.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130449.
PR 23-APR-1999; 9905-0130510.
PR 23-APR-1999; 9905-0130891.
PR 28-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132048.
PR 30-APR-1999; 9905-0132407.
PR 04-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 07-MAY-1999; 9905-0132487.
PR 11-MAY-1999; 9905-0134256.
PR 11-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.

PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144633.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144885.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145291.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157153.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 10.3%; Score 6; DB 21; Length 169;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 KKYAYI 51
|||||
DB 21 KKYAYI 26

RESULT 14
AAU05778
ID AAU05778 standard; Protein; 178 AA.
XX
AC AAU05778;
XX
DF 24-OCT-2001 (first entry)
DE
XX Soybean Invertase Inhibitor #3.
XX
KW Soybean; Invertase Inhibitor; gene therapy; environmental stress;
KW kernel development; antisense.
XX
OS Glycine max.
XX
PN WO200158939-A2.
XX
PD 16-Aug-2001.
XX
XX 12-FEB-2001; 2001WO-US04492.
PF
XX 10-FEB-2000; 2000US-0181509.
PR
XX

PA (PION-) PIONEER HI-BRED INT INC.
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Helentjaris T, Bate NJ, Allen SM;
XX
DR WPI; 2001-502706/55.
DR N-PSDB; AAS11363.
XX
PT An isolated polypeptide when recombinantly expressed in a plant is
PT useful for modulating invertase activity and increased in yield in the
PT plant -
XX
PS Claim 1; Page 71; 83pp; English.
XX
XX The invention relates to novel plant invertase inhibitors or a
CC yeast invertase which, when recombinantly expressed in a plant, can
CC modulate invertase activity and increase yield in the plant. Chimeric
CC invertase inhibitors are useful for modulating invertase activity and
CC increasing yield in a plant (especially crop species) when used to
CC transform the plant and are also useful for modulating kernel development
CC and protecting plants against the harmful/detrimental effects of stress
CC and adverse environmental conditions. Yeast invertase is less sensitive
CC to invertase inhibitors therefore is an attractive option to supplement
CC invertase activity in a plant using gene therapy. The novel
CC invertase inhibitor nucleic acids may be used in their antisense form.
CC The present sequence represents a soybean invertase inhibitor.
CC
XX
SQ Sequence 178 AA;
Query Match 10.3%; Score 6; DB 22; Length 178;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 KYKTVL 9
|||||
DB 107 KYKTVL 112
RESULT 15
AAU25554
ID AAU25554 standard; Protein; 180 AA.
XX
AC AAU25554;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human G Protein-Coupled Receptor (GPCR) polypeptide #1.
XX
KW Human; G-protein coupled receptor; GPCR; mental disorder; schizophrenia;
KW attention deficit disorder; anxiety; depression; bipolar disorder;
KW neurological disorder; Huntington's disease; dementia; obesity; anorexia;
KW metabolic disorder; Parkinson's disease; Tourette's syndrome; thrombosis;
KW type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer;
KW cardiomyopathy; atherosclerosis; human immunodeficiency virus; HIV;
KW viral infection; immunostimulant; neuroleptic; nootropic; tranquiliser;
KW antidepressant; anorectic; gene therapy.
XX
XX Homo sapiens.
XX
PN WO200162797-A2.
PD 30-AUG-2001.
XX
XX 23-FEB-2001; 2001MO-US05676.
PF
XX 23-FEB-2000; 2000US-0184247.
PR 23-FEB-2000; 2000US-0184303.
PR 23-FEB-2000; 2000US-0184304.
PR 23-FEB-2000; 2000US-0184305.
PR 23-FEB-2000; 2000US-0184397.
PR 03-MAR-2000; 2000US-0186457.
PR 03-MAR-2000; 2000US-0186810.
PR 09-MAR-2000; 2000US-0188064.

PR 13-MAR-2000; 2000US-0188880.
PR 03-APR-2000; 2000US-0194344.
PR 23-JUN-2000; 2000US-0213861.
PR 11-JUL-2000; 2000US-0217369.
PR 11-JUL-2000; 2000US-0217370.
PR 14-JUL-2000; 2000US-0218337.
PR 20-JUL-2000; 2000US-0218492.
XX
XX (PHAA) PHARMACIA & UPJOHN CO.
XX
XX Vogel G, Wood LS, Parodi LA, Lind P;
XX WPI; 2001-570628/64.
XX N-PSDB; AAS42806.
XX
XX New isolated nucleic acid encoding a new G-protein coupled receptor
XX polypeptide for detecting receptor modulators that can treat mental
XX disorders, such as schizophrenia, anxiety, depression, or obesity -
XX
XX Claim 35; Page 70; 279pp; English.
XX
XX Sequences AAU25554-AAU25616 represent human G-protein coupled receptor
XX (GPCR) polypeptides of the invention. The proteins and their associated
XX DNA sequences can be used to identify compounds which bind to GPCR
XX polypeptides and in screening for compounds that modulate GPCR activity.
XX By screening a human subject for the presence of mutations in GPCR DNA, a
XX GPCR-related disorder or a genetic predisposition can be diagnosed. The
XX sequences can also be used for treatment and prevention of mental
XX disorders such as schizophrenia, attention deficit disorder, anxiety,
XX depression, dementia and bipolar disorder, neurological disorders such as
XX Huntington's disease, Parkinson's disease and Tourette's syndrome,
XX metabolic disorders such as obesity, anorexia and type 2 diabetes,
XX cardiovascular disorders such as thrombosis, myocardial infarction,
XX cardiomyopathy and atherosclerosis, viral infections caused by HIV and
XX cancers.
XX
XX
SQ Sequence 180 AA;
Query Match 10.3%; Score 6; DB 22; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 41 YNLFTK 46
|||||
DB 136 YNLFTK 141
RESULT 16
AAB96487
ID AAB96487 standard; Protein; 186 AA.
XX
AC AAB96487;
XX
DT 29-OCT-2001 (first entry)
XX
DE Putative P. abyssal nucleotidyltransferase.
XX
XX Hyperthermophilic archaeon; hyperthermophilic protein.
XX
XX Pyrococcus abyssi.
XX
XX FR2792651-A1.
XX
XX 27-OCT-2000.
XX
XX 21-APR-1999; 99FR-0005034.
XX
XX 21-APR-1999; 99FR-0005034.
XX
XX (CNRS) CNRS CENT NAT RECH SCI.
XX (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
XX Forterre P, Thiery JC, Prieur D, Dietrich J, Lecompte O;

PI Querellou J, Weissenbach J, Saurin W, Hellig R;
XX WPI; 2001-126236/14.
XX
PT New nucleotide sequences isolated from *Pyrococcus abyssi* encode
XX proteins useful in industry -
XX
PS Claim 7; Page 1196; 1657pp; French.
XX
CC The present invention relates to the genomic sequence of *Pyrococcus*
CC *abyssi* (see AAF6431 and AAH1223-77) and *P. abyssi* proteins. *P. abyssi* is
CC a hyperthermophilic archaean, which is isolated from deep-sea
CC hydrothermal vents. The present sequence is one such *P. abyssi* protein.
CC The proteins of the present invention have various potential industrial
CC uses, since the proteins are stable at very high temperatures, some up to
CC 110 degrees centigrade.
CC Note: This patent is in the same patent family as WO200065062, which
CC contains additional sequences as shown in AAB9132-AAB9143,
CC AAH75903-AAH75920 and AAC66436.
XX
SQ Sequence 186 AA;

Query Match 10.3%; Score 6; DB 22; Length 186;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 EYKEI 58
Db 152 EYKEI 157

RESULT 17
AAU17158
ID AAU17158 standard; Protein; 188 AA.
XX
AC AAU17158;
XX
DT 07-NOV-2001 (first entry)
XX
DE Novel signal transduction pathway protein, Seq ID 723.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS;
KW acquired immune deficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200154733-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01312.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-019874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225269.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241321.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.

PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-465460/50.

DR N-PSDB; AAS27075.

PT Novel polypeptides useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cancers, immune
 PT disorders and neuronal disorders -

XX Claim 1; SEQ ID No 723; 880pp; English.

XX The invention relates to novel isolated polypeptides (I), and
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
 CC diagnosing, preventing and treating diseases including immune system
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 CC transplant rejections and graft versus host disease, infectious diseases
 CC (e.g. hepatitis C), bleeding disorders, hemoglobin abnormalities and
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative
 CC disorders, primary haematopoietic disorders, hyperproliferative

CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
 CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
 CC Addison's disease), reproductive system disorders, gastrointestinal
 CC disorder (inflammatory disorders), liver disorders (cirrhosis),
 CC as stimulators of B-cell responsiveness to pathogens, activators of
 CC T-cells, to induce higher affinity antibodies, and as a means to induce
 CC tumour proliferation in pathological e.g. acquired immune deficiency
 CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
 CC pathway protein, amino acid sequences of the invention.

Query Match 10.3%; Score 6; DB 22; Length 188;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 KVVYSY 38~
 |||||
 DB 92 KVVYSY 97

RESULT 18
 AAB33164
 ID AAB33164 standard; Protein; 195 AA.

XX AAB33164;

PT 25-JAN-2001 (first entry)

XX Eucalyptus grandis transcription factor protein sequence #353.

XX Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
 XX poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
 XX basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
 XX homeodomain zipper; LIM domain; AP2; EREBS; zinc finger domain;
 XX type 2 Cys2His2; CCAAT box element; MYB.

XX Eucalyptus grandis.

XX WO200053724-A2.

XX 14-SEP-2000.

XX 09-MAR-2000; 2000MO-US06112.

XX 11-MAR-1999; 99US-0266513.

XX 18-AUG-1999; 99US-0149485.

XX (GENE-) GENESIS RES & DEV CORP LTD.
 XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX Wood M, McGrath A, Shenk MA, Glenn M;

XX WPI; 2000-579369/54.

PT New isolated polynucleotide encoding a plant transcription factor for
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
 PT having modified gene expression or modified activity of a polypeptide

XX Claim 8; Page 662; 747pp; English.

XX The present invention relates to novel plant transcription factors from
 CC Eucalyptus grandis or Pinus radiata. The present sequence is one such
 CC transcription factor. The transcription factor may be used to produce a
 CC plant having modified gene expression such as a woody plant e.g. a
 CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or
 CC to modify the activity of a polypeptide in a plant. The transcription
 CC factors of the present invention are members from the following families
 CC of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic

CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain
CC zipper, LIM domain, AP2 and EREBs, zinc finger domains of type 2
CC Cys2His2, CCAAT box elements and MYB.
XX
SQ Sequence 195 AA;

Query Match 10.3%; Score 6; DB 21; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ERKXT 7
Db 99 ERKXT 104

RESULT 19
AA08657
ID AA08657 standard; Protein; 196 AA.

XX
XX
AC AA08657;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 6281.

XX
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

XX
XX
PN EP1033405-A2.

XX
PD 06-SEP-2000.

PE 25-FEB-2000; 2000EP-0301439.

XX
XX
PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 07-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138340.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 17-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139452.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 24-JUN-1999; 99US-0140354.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142380.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144864.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151348.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158569.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

10.3%; Score 6; DB 21; Length 196;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
33 KVVYSY 38
|||||
117 KVVYSY 122
RESULT 20
ABBI7135
ID ABBI7135 standard; Protein; 223 AA.
23-JAN-2002 (first entry)
Human nervous system related polypeptide SEQ ID NO 5792.
Human; nootropic; neuroprotective; cyostatic; dermatological; virucide;
Immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
antiParkinsonian; antisticking; antianaemic; antiarthritic; cancer;
antipneumatic; hepatoprotective; cerebroprotective; antiinflammatory;
antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
antiparasitic; cardiatic; immune disorder; cardiovascular disorder;
neurological disease; infection; nephrotropic; gene therapy; vaccine.
Homo sapiens.
WO200159063-A2.
16-AUG-2001.
17-JAN-2001; 2001WO-US01334.
31-JAN-2000; 2000US-0179065.
04-FEB-2000; 2000US-0180628.
24-FEB-2000; 2000US-0184664.
02-MAR-2000; 2000US-0186350.
16-MAR-2000; 2000US-0189874.
17-MAR-2000; 2000US-0190076.
18-APR-2000; 2000US-0198123.
19-MAY-2000; 2000US-0205515.
07-JUN-2000; 2000US-0209467.
28-JUN-2000; 2000US-0214886.
30-JUN-2000; 2000US-0215135.
07-JUL-2000; 2000US-0216647.
07-JUL-2000; 2000US-0216880.
11-JUL-2000; 2000US-0217487.
11-JUL-2000; 2000US-0217496.
14-JUL-2000; 2000US-0218290.
26-JUL-2000; 2000US-0220963.
14-AUG-2000; 2000US-0220964.
14-AUG-2000; 2000US-0224518.
14-AUG-2000; 2000US-0224519.
14-AUG-2000; 2000US-0225213.
14-AUG-2000; 2000US-0225214.
14-AUG-2000; 2000US-0225266.
14-AUG-2000; 2000US-0225267.
14-AUG-2000; 2000US-0225268.
14-AUG-2000; 2000US-0225270.
14-AUG-2000; 2000US-0225447.
14-AUG-2000; 2000US-0225757.
14-AUG-2000; 2000US-0225758.
14-AUG-2000; 2000US-0225759.
18-AUG-2000; 2000US-0226279.
22-AUG-2000; 2000US-0226681.
22-AUG-2000; 2000US-0226688.
22-AUG-2000; 2000US-0227182.
23-AUG-2000; 2000US-0227009.
30-AUG-2000; 2000US-0228924.
01-SEP-2000; 2000US-0229387.
01-SEP-2000; 2000US-0229343.
01-SEP-2000; 2000US-0229344.

[illegible]

	17-NOV-2000;	2000US-0249215.	
RR	17-NOV-2000;	2000US-0249216.	
PR	17-NOV-2000;	2000US-0249217.	
PR	17-NOV-2000;	2000US-0249218.	
PR	17-NOV-2000;	2000US-0249219.	
PR	17-NOV-2000;	2000US-0249220.	
PR	17-NOV-2000;	2000US-0249221.	
PR	17-NOV-2000;	2000US-0249222.	
PR	17-NOV-2000;	2000US-0249223.	
PR	17-NOV-2000;	2000US-0249224.	
PR	17-NOV-2000;	2000US-0249225.	
PR	17-NOV-2000;	2000US-0249226.	
PR	17-NOV-2000;	2000US-0249227.	
PR	17-NOV-2000;	2000US-0249228.	
PR	17-NOV-2000;	2000US-0249229.	
PR	17-NOV-2000;	2000US-0249230.	
PR	17-NOV-2000;	2000US-0250391.	
PR	01-DEC-2000;	2000US-0251160.	
PR	05-DEC-2000;	2000US-0251030.	
PR	05-DEC-2000;	2000US-0251030.	
PR	05-DEC-2000;	2000US-0256719.	
PR	06-DEC-2000;	2000US-0251479.	
PR	08-DEC-2000;	2000US-0251856.	
PR	08-DEC-2000;	2000US-0251856.	
PR	08-DEC-2000;	2000US-0251869.	
PR	08-DEC-2000;	2000US-0251989.	
PR	08-DEC-2000;	2000US-0251990.	
PR	11-DEC-2000;	2000US-0254097.	
PR	05-JAN-2001;	2001US-0259678.	
XX	(HUMA-) HUMAN GENOME SCI INC.		
PA	Rosen CA,	Barash SC,	Ruben SM;
PI	WPI:	2001-541565/60.	
XX	N-PSDB; ABAI3461.		
DR	Nucleic acids encoding 3224 human nervous system antigen polypeptides,		
PT	useful for preventing, diagnosing and/or treating nervous system		
PT	cancers and metastases -		
XX	Claim 11; SEQ ID NO 5792; 1701bp + Sequence Listing; English.		
PS	The invention relates to novel genes (ABAI1004-ABA21534) and proteins		
XX	(ABAI1678-ABAI8001) useful for preventing, treating or ameliorating		
CC	medical conditions e.g., by protein or gene therapy. The genes are		
CC	isolated from a range of human tissues disclosed in the specification.		
CC	The nucleic acids, proteins, antibodies and (ant)agonists are useful		
CC	in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast		
CC	and ovarian cancer and other cancers of the adrenal gland, bone, bone		
CC	marrow, breast, gastrointestinal tract, liver, lung, or urogenital;		
CC	(b) immune disorders e.g. Addison's disease, allergies, autoimmune		
CC	haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's		
CC	disease, multiple sclerosis, Rheumatoid arthritis and ulcerative		
CC	colitis; (c) cardiovascular disorders such as myocardial ischemias;		
CC	(d) wound healing; (e) neurological diseases e.g. cerebral anoxia and		
CC	epilepsy; and (f) infectious diseases such as viral, bacterial, fungal		
CC	and parasitic infections.		
CC	Note: The sequence data for this patent did not form part of the		
CC	printed specification, but was obtained in electronic format directly		
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences.		
XX	Sequence 223 AA;		
SQ	Query Match 10.3%; Score 6; DB 22; Length 223; Best Local Similarity 100.0%; Pred. No. 1.5e+02; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	32 AKVYS 37 		
Db	81 AKVYS 86		
NC	RESULT 21		
AAU19799	AAU19799 standard; Protein; 223 AA.		
XX	AAU19799;		

XX 06-DEC-2001 (first entry)
DT Human novel extracellular matrix protein, Seq ID No 449.
XX
DE
XX Human; secreted extracellular matrix protein; immunomodulatory;
KW Anti-HIV; antineoplastic; antirheumatic; antisclerotic; cardiant; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;
KW antileishmaniasis; immune/autoimmune disease; HIV infection; anaemia;
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;
KW Sezary syndrome; Gaucher's disease; neurological diseases;
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
KW cardiac arrest; tachycardia; angina; infection; corneal infections;
KW wound healing; immunogen; gene therapy; antisense; food additive.
XX
OS Homo sapiens.
XX WO200155368-A1.
XX
XX 02-AUG-2001.
PD
XX
PF 17-JAN-2001; 2001WO-US01348.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.

PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI: 2001-465572/50.
 DR N-PSDB; AAS313370.
 XX
 XX Nucleic acid molecules encoding human secreted extracellular matrix
 PT proteins, used in preventing, treating or ameliorating a disorder, e.g.
 PT Alzheimer's and Parkinson's diseases and cancers -
 XX
 PS Claim 11; SEQ ID NO 449; 577pp; English.
 XX
 CC The invention relates to isolated nucleic acid molecules encoding
 CC novel human secreted extracellular matrix proteins (SPs). The
 CC polynucleotides and proteins are used to prevent, treat a medical
 CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
 CC chickens or sheep. For example, disorders associated with decreased
 CC expression of SPs. The SP polynucleotide or a vector expressing them may
 CC be administered to treat diseases by gene therapy. Antisense molecules
 CC may be administered to down regulate expression of SPs by binding with
 CC the cells own genes and preventing their expression. The polynucleotides
 CC may also be used as DNA probes in diagnostic assays. The SPs may also be
 CC used as antigens to produce antibodies and to identify modulators
 CC (agonists and antagonists) of the SPs. The anti-(SP) antibodies and
 CC antagonists may also be used to down regulate expression and activity of
 CC SP and as diagnostic agents for detecting the presence of SPs in samples.
 CC The disorders include for example: immune/autoimmune diseases (e.g. HIV
 CC (human immunodeficiency virus) infections, anemia, rheumatoid arthritis
 CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
 CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
 CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
 CC Parkinson's disease) cardio-/cerebrovascular disorders (e.g. cardiac
 CC arrest, tachycardia and angina), infections caused by bacteria, viruses
 CC and fungi and ocular disorders (e.g. corneal infections). Other uses
 CC include wound healing, maintenance of organs before transplantation,
 Query Match 10.3%; Score 6; DB 22; Length 223;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
 KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;
 KW type 2 Cys2His2; CCAAT box element; MYB.
 XX
 OS Eucalyptus grandis.
 XX
 PN W0200053724-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06112.
 XX
 PR 11-MAR-1999; 99US-0266513.
 PR 18-AUG-1999; 99US-0149485.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 XX
 PI Wood M, McGrath A, Shenk MA, Glenn M;
 XX
 DR WPI: 2000-579369/54.
 XX
 PS
 PS
 PS Claim 8; Page 203; 747pp; English.
 XX
 CC The present invention relates to novel plant transcription factors from
 CC Eucalyptus grandis or Pinus radiata. The present sequence is one such
 CC transcription factor. The transcription factor may be used to produce a
 CC plant having modified gene expression such as a woody plant e.g. a
 CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or
 CC to modify the activity of a polypeptide in a plant. The transcription
 CC factors of the present invention are members from the following families
 CC of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic
 CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain
 CC zipper, LIM domain, AP2 and EREBs, zinc finger domains of type 2
 CC Cys2His2, CCAAT box elements and MYB.
 CC
 XX Sequence 233 AA;
 XX
 Query Match 10.3%; Score 6; DB 21; Length 233;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 AKVYYS 37
 |||||
 DB 81 AKVYYS 86
 RESULT 22
 AAB32552
 ID AAB32552 standard; Protein; 233 AA.
 XX
 AC AAB32552;
 XX
 DT 25-JAN-2001 (first entry)
 XX
 DE Eucalyptus grandis transcription factor protein sequence #10.
 XX
 KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
 KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
 KW

RESULT 23
 AAB18242
 ID AAB18242 standard; Protein; 235 AA.
 XX
 AC AAB18242;
 XX
 DT 07-NOV-2000 (first entry)
 XX
 DE Plasmodium falciparum chromosome 2 related protein SEQ ID NO:99.
 XX
 KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
 KW antimalarial; malaria; protozoacide; infection; insecticide.
 XX
 OS Plasmodium falciparum.
 XX
 PN W0200025728-A2.
 XX
 PD 11-MAY-2000.
 XX
 PF 05-NOV-1999; 99WO-US26796.
 XX

PF 22-MAR-2000; 2000WO-0507506.
XX
PR 26-MAR-1999; 99US-0126505.
PR 17-DEC-1999; 99US-0172412.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM, Komatsoulis G;
XX WPI; 2000-594649/56.
DR
XX Forty-nine polynucleotide sequences, and their encoded secreted
PT polypeptides, used in the treatment and diagnosis of cancers,
PT autoimmune disorders, and skin disorders -
XX
PS Disclosure; Pages 393-394; 413pp; English.
XX
CC The invention relates to the isolation of genes AAC73865-C73913 encoding
CC the human secreted proteins AAB39093-839141. This sequence was used as a
CC query sequence for doing BLASTX searches to determine homologous
CC sequences to the protein isolated in the present invention. The genes
CC and proteins are useful for preventing, ameliorating or treating medical
CC conditions, e.g. by protein or gene therapy. The genes are isolated
CC from a range of human tissues disclosed in the specification. The nucleic
CC acids, proteins, antibodies and (ant)agonists are useful in the
CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b)
CC immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischemias; (d)
CC wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
XX
SQ Sequence 241 AA;
XX
Query Match 10.38; Score 6; DB 21; Length 241;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 41 YNLFETK 46
Db 232 YNLFETK 237
XX
RESULT 26
AAG05491
ID AAG05491 standard; Protein; 247 AA.
XX
AC AAG05491;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 1916.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
PD
XX 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155569.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

10.3%; Score 6; DB 21; Length 247;

Best Local Similarity 100.0%; Pred. No. 1.7e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0;

53 EYKEI 58

157 EYKEI 162

RESULT 27

AAG05490 standard; Protein; 248 AA.

AAG05490;

17-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 1915.

Protein identification: signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-0301439.

25-FEB-1999; 99US-0121825.

05-MAR-1999; 99US-0123180.

09-MAR-1999; 99US-0123548.

23-MAR-1999; 99US-0125788.

25-MAR-1999; 99US-0126264.

29-MAR-1999; 99US-0126785.

01-APR-1999; 99US-0127462.

06-APR-1999; 99US-0128234.

08-APR-1999; 99US-0128714.

16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135253.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 12-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match Best Local Similarity 10.3%; Score 6; DB 21; Length 248;
Pred. No. 1.7e+02; Mismatches 0; Indels 0; Gaps 0;

OY 53 EYKEI 58
DB 158 EYKEI 163

RESULT 28

ABB64578 ID ABB64578 standard; Protein; 251 AA.

XX ABB64578;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 20526.

KW Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL08681.

PT New isolated nucleic acid detection reagent for detecting 1000 or more

PS genes from Drosophila and for elucidating cell signalling and cell-cell

CC interactions -
The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB57737-AB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 251 AA;

Query Match Best Local Similarity 10.3%; Score 6; DB 22; Length 251;
Pred. No. 1.7e+02; Mismatches 0; Indels 0; Gaps 0;

OY 19 FPHQIS 24
DB 39 FPHQIS 44

RESULT 29

ABB64575 ID ABB64575 standard; Protein; 262 AA.

XX ABB64575;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 57510.

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145081.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156459.
PR 28-SEP-1999; 99US-0156548.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 10.3%; Score 6; DB 21; Length 262;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 EYKEI 58
|||||
DB 42 EYKEI 47

RESULT 30
ABB58183

ID ABB58183 standard; Protein; 262 AA.

AC ABB58183;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 1341.

KW Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL02286.

PS Disclosure; SEQ ID NO 1341; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

CC Sequence 262 AA;

Query Match 10.3%; Score 6; DB 22; Length 262;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 FPHOIS 24
|||||
DB 48 FPHOIS 53

RESULT 31
ABB63475
ID ABB63475 standard; Protein; 272 AA.

AC ABB63475;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 17217.

KW Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL07578.

PS Disclosure; SEQ ID NO 17217; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

CC Sequence 272 AA;

Query Match 10.3%; Score 6; DB 22; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FHOISM 25
|||||
DB 48 FHOISM 53

RESULT 32

ABB67569
ID ABB67569 standard; Protein; 288 AA.

AC ABB67569;

DT 29-MAY-2001 (first entry)

DE Amino acid sequence of a human hydrolytic enzyme HYENZ1.

KW Human; hydrolytic enzyme; HYENZ; neurological disorder: cancer;
immune system disorder; genetic disorder; cell proliferation disorder;

KW	epilepsy	ischemic cerebrovascular disease; stroke; Pick's disease;
KW	Huntington's disease; dementia; Parkinson's disease; multiple sclerosis;	
KW	viral meningitis; Creutzfeldt-Jakob disease; neurofibromatosis;	
KW	cerebral palsy; autonomic nervous system disorder; mental disorder;	
KW	cerebral nerve disorder; peripheral nervous system disorder;	
KW	immune system disorder; osteoarthritis; gene therapy.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Modified-site	12
FT	Modified-site	/note= "potential phosphorylation site"
FT	Modified-site	95
FT	Modified-site	/note= "potential glycosylation site"
FT	Modified-site	136
FT	Modified-site	/note= "potential phosphorylation site"
FT	Modified-site	172
FT	Modified-site	/note= "potential phosphorylation site"
FT	Modified-site	183
FT	Modified-site	/note= "potential phosphorylation site"
FT	Modified-site	196
FT	Modified-site	/note= "potential phosphorylation site"
FT	Modified-site	215
FT	Modified-site	/note= "potential phosphorylation site"
FT	Modified-site	237
FT	Modified-site	/note= "potential phosphorylation site"
FT	Modified-site	276
FT	Modified-site	/note= "potential phosphorylation site"
FT	Modified-site	278
FT	Modified-site	/note= "potential phosphorylation site"
XX		
PN	MO200116334-A2.	
XX		
PD	08-MAR-2001.	
XX		
PF	31-AUG-2000; 2000WO-US24107.	
XX		
PR	01-SEP-1999; 99US-0151819.	
XX		
PA	(INCY-) INCYTE GENOMICS INC.	
XX		
P1	Yue H, Hillman JL, Tang YT, Baughn MR, Lu DAM, Azimzai Y;	
XX		
DR	WPI; 2001-235111/24.	
DR	N-PSDB; AAF55425.	
XX		
PT	Novel human hydrolytic enzymes useful for diagnosing, treating, or	
PT	preventing disorders associated with abnormal expression of HYENZ, cell	
PT	proliferative disorders, neurological disorders and immune system	
PT	disorders -	
XX		
PS	Claim 1; Page 89-90; 113pp; English.	
XX		
CC	The present sequence represents a human hydrolytic enzyme (HYENZ). The	
CC	specification describes HYENZ-1 to HYENZ-14. HYENZ polypeptides and	
CC	polynucleotides are useful in the diagnosis, prevention and treatment of	
CC	neurological disorders, immune system disorders, genetic disorders, and	
CC	cell proliferation disorders including cancer. They are useful for	
CC	treating epilepsy, ischemic cerebrovascular disease, stroke, Pick's	
CC	disease, Huntington's disease, dementia, Parkinson's disease, multiple	
CC	sclerosis, viral meningitis, Creutzfeldt-Jakob disease,	
CC	neurofibromatosis, cerebral palsy, autonomic nervous system disorder,	
CC	cranial nerve disorder, peripheral nervous system disorder, mental	
CC	disorders, immune system disorders, osteoarthritis, and genetic	
CC	disorders. HYENZ polynucleotides are useful for somatic or germline	
CC	gene therapy for treating the disorders.	
XX		
SO	Sequence	288 AA;
XX		
Query Match	10.3%; Score 6; DB 22; Length 288;	
Best Local Similarity	100.0%; Pred. No. 1.9e+02;	
Matches	6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	

QY	26	FEDLYD	31
DB	174	FEDLYD	179
RESULT	N	33	
ID	ABG30285		
	ABG30285 standard; Protein; 289 AA.		
XX			
AC	ABG30285;		
XX			
DT	18-FEB-2002 (first entry)		
DE	Novel human diagnostic protein #30276.		
XX			
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;		
KW	food supplement; medical imaging; diagnostic; genetic disorder.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200175067-A2.		
PD	11-OCT-2001.		
XX			
PF	30-MAR-2001; 2001WO-US08631.		
PR	31-MAR-2000; 2000US-0540217.		
PR	23-AUG-2000; 2000US-0649167.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Drimac RT, Liu C, Tang YT;		
DR	WI: 2001-639362/73.		
DR	N-PSDB; AAS94472.		
XX			
PT	New isolated polynucleotide and encoded polypeptides, useful in		
PT	diagnostics, forensics, gene mapping, identification of mutations		
PT	responsible for genetic disorders or other traits and to assess		
PT	biodiversity -		
XX			
PS	Claim 20; SEQ ID NO 60644; 103pp; English.		
XX			
CC	The invention relates to isolated polynucleotide (I) and		
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,		
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome		
CC	and gene mapping, and in recombinant production of (II). The		
CC	polynucleotides are also used in diagnostics as expressed sequence tags		
CC	for identifying expressed genes. (I) is useful in gene therapy techniques		
CC	to restore normal activity of (II) or to treat disease states involving		
CC	((II)). ((II) is useful for generating antibodies against it, detecting or		
CC	quantitating a polypeptide in tissue, as molecular weight markers and as		
CC	a food supplement. ((II) and its binding partners are useful in medical		
CC	imaging of sites expressing ((I)). ((I) and ((II) are useful for treating		
CC	disorders involving aberrant protein expression or biological activity.		
CC	The polypeptide and polynucleotide sequences have applications in		
CC	diagnostics, forensics, gene mapping, identification of mutations		
CC	responsible for genetic disorders or other traits to assess biodiversity		
CC	and to produce other types of data and products dependent on DNA and		
CC	amino acid sequences. ABG00010-ABG30377 represent novel human		
CC	diagnostic amino acid sequences of the invention.		
CC	Note: The sequence data for this patent did not appear in the printed		
CC	specification, but was obtained in electronic format directly from WIPO		
CC	at ftp.wipo.int/pub/published_pct_sequences.		
XX			
SQ	Sequence	289	AA;
Query Match	10.3%;	Score 6;	DB 22; Length 289;
Best Local Similarity	100.0%;	Pred. No. 1.9e+02;	
Matches	6;	Conservative	0; Mismatches 0; Indels 0; Gaps 0;
XY	27	EDLYDA	32

DB 26 EDLYDA 31

RESULT 34
ID AAB53344 standard; Protein: 293 AA.

XX AAB53344;
AC AAB53344;
XX 09-MAR-2001 (first entry)
DT 09-MAR-2001 (first entry)
XX Human colon cancer antigen protein sequence SEQ ID NO:884.
DE Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX Identification; cytostatic; cardioprotective; neuroprotective; vulnerary;
KW immunomodulatory; muscular; gynaecological; gastrointestinal;
KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW neural disorder; immune system disorder; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; renal disorder;
KW infectious disease; cardiovascular disorder.

XX Homo sapiens.
OS WO200053531-A1.
PN 21-SEP-2000.
PD 08-MAR-2000; 2000MO-US05883.
PE 12-MAR-1999; 99US-0124270.
PR (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Ruben SM;
PI WPI: 2000-587534/55.
DR N-PSDB; AAC98101.

XX Colon cancer associated gene sequences, referred to as colon cancer
PT antigens, useful for the treatment, prevention, and diagnosis of colon
PT disorders such as colon cancer -
PS Claim 11: Page 1437-1438; 2104pp; English.

XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC human colon cancer antigens can have cytostatic, cardioprotective, muscular;
CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC vulnerary, nephrotropic, antiinfective and antibacterial activities, and
CC can be used in gene therapy. The colon cancer antigen polynucleotides,
CC proteins and antibodies to the proteins are useful for the prevention,
CC treatment and diagnosis of colon disorders, such as colon cancer. The
CC polynucleotides may be used in diagnostics and research, such as for
CC chromosome identification, and as hybridisation probes. The proteins
CC may also be used to prevent diseases such as neural disorders, immune
CC system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, wounds, renal disorders, infectious
CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and
CC AAB54007 represent sequences used in the exemplification of the present
CC invention.

XX Sequence 293 AA;
SQ

Query Match 10.3%; Score 6; DB 21; Length 293;
Best Local Similarity 100.0%; Pred.No.1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 41 YNLFTR 46
DB 283 YNLFTR 288

RESULT 35

AAG41745
ID AAG41745 standard; Protein: 294 AA.
XX AAG41745;
AC AAG41745;
XX 18-OCT-2000 (first entry)
DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 51976.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 51976.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS EPI033405-A2.
PN 06-SEP-2000.
PD 25-FEB-2000; 2000EP-0301439.
PE 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PA 09-MAR-1999; 99US-0123548.
PI 23-MAR-1999; 99US-0125788.
DR 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 04-MAY-1999; 99US-0132048.
XX 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132853.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141847.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143547.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 10.3%; Score 6; DB 21; Length 294;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 34 VVSYX 39
|||||
Db 228 VVSYX 233

RESULT 36
AAG41744
ID AAG41744 standard; Protein: 296 AA.

XX AAG41744;
AC
XX
DT 18-OCT-2000 (first entry)
XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 51975.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX
XX EF1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.

PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152353.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155139.
 PR 24-SEP-1999; 99US-0155486.
 PR 28-SEP-1999; 99US-0156559.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159283.
 PR 13-OCT-1999; 99US-0159284.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161982.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 10.3%; Score 6; DB 21; Length 296;
 Best Local Similarity 100.0%; Pred. No. 2e+02; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0;

QY 34 VVSYVY 39
 |||||
 Db 230 VVSYVY 235

RESULT 37
 AAY73880
 ID AAY73880 standard; Protein; 298 AA.
 XX
 AC AAY73880;

XX 14-MAR-2000 (first entry)
 DE Human prostate tumor EST fragment derived protein #67.
 XX
 KW Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;
 KW treatment.
 XX
 OS Homo sapiens.
 XX

PN DE19820190-A1.
 XX
 PD 04-NOV-1999.
 XX
 PF 28-APR-1998; 98DE-1020190.
 XX
 PR 28-APR-1998; 98DE-1020190.
 XX
 PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX
 PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
 XX
 DR WPI: 1999-621386/54.
 DR N-PSDB; AA252880.
 XX
 PT New human nucleic acid sequences from pancreatic tumors, and related
 PT proteins
 XX
 PS Claim 23; Page 336; 502pp; German.
 XX
 CC This invention describes novel polypeptides and their encoding nucleic
 CC acids derived from human pancreatic tumor tissue which have cytostatic
 CC activity. The sequences are also useful in producing pharmaceutical
 CC compositions for treatment of pancreatic tumors. AAY73814-Y74252
 CC represent protein fragments encoded by the human pancreatic tumor cDNA
 CC library derived expressed sequence tag (EST) sequences represented in
 CC AA252858-253014.
 XX
 SQ Sequence 298 AA;

Query Match 10.3%; Score 6; DB 20; Length 298;
 Best Local Similarity 100.0%; Pred. No. 2e+02; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0;

QY 26 FEDLYD 31
 |||||
 Db 184 FEDLYD 189

RESULT 38
 AAB43779
 ID AAB43779 standard; Protein; 298 AA.
 XX
 AC AAB43779;

DE 08-FEB-2001 (first entry)

XX Human cancer associated protein sequence SPQ ID NO:1224.

KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnerrary; immunomodulator;
 KW antidiabetic; antitumoral; antitumoral; antitumoral; antitumoral;
 KW antitumoral; antitumoral; antitumoral; antitumoral; antitumoral;
 KW dermatological; neuroproliferative; thrombolytic; coagulant; nootropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.

XX Homo sapiens.

XX WO20005350-A1.

XX 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05882.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

XX WPI; 2000-587533/55.
DR N-PSDB; AAC77988.
XX
PT Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer -
XX
PS Claim 11; Page 1856-1857; 2352pp; English.
XX
CC AAC77607 to AAC78448 encode the human cancer associated proteins given
CC in ABA43398 to ABA44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnery; immunomodulator;
CC antidiabetic; antiallergic; antineumatic; antiarthritic;
CC antiinflammatory; antihypertensive; antiallergic; antibacterial;
CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC neotropic; vasotropic; antiproliferative and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and ABA44240 represent sequences used in the exemplification of
CC the present invention.
XX
SQ Sequence 298 AA;

Query Match 10.3%; Score 6; DB 21; Length 298;
Best Local Similarity 100.0%; Pred. No. 2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 26 FEDLYD 31
Db 184 FEDLYD 189

RESULT 39
AAU23487
ID AAU23487 standard; Protein; 298 AA.
XX
AC AAU23487;
XX
DT 17-DEC-2001 (first entry)
XX
XX Novel human enzyme polypeptide #573.
DE
XX
XX Human: oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
XX ligase; hyperproliferative disorder; immunodeficiency disorder;
XX autoimmune disorder; neurological disorder; metabolic disorder;
XX inflammatory disorder; cardiovascular disorder; reproductive disorder;
XX blood-related disorder; infectious disorder; cystostatic; anti arthritic;
XX nephrotropic; anticoagulant.
OS Homo sapiens.
XX
XX WO200155301-A2.
PN
XX
PD 02-AUG-2001.
XX
XX
PF 17-JAN-2001; 2001WO-US01239.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 14-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228824.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.

PR	20-OCT-2000	2000US-0240960
PR	20-OCT-2000	2000US-0241221
PR	20-OCT-2000	2000US-0241785
PR	20-OCT-2000	2000US-0241786
PR	20-OCT-2000	2000US-0241787
PR	20-OCT-2000	2000US-0241808
PR	20-OCT-2000	2000US-0241809
PR	20-OCT-2000	2000US-0241826
PR	01-NOV-2000	2000US-0244617
PR	08-NOV-2000	2000US-0246474
PR	08-NOV-2000	2000US-0246475
PR	08-NOV-2000	2000US-0246476
PR	08-NOV-2000	2000US-0246477
PR	08-NOV-2000	2000US-0246478
PR	08-NOV-2000	2000US-0246523
PR	08-NOV-2000	2000US-0246524
PR	08-NOV-2000	2000US-0246525
PR	08-NOV-2000	2000US-0246526
PR	08-NOV-2000	2000US-0246527
PR	08-NOV-2000	2000US-0246528
PR	08-NOV-2000	2000US-0246532
PR	08-NOV-2000	2000US-0246609
PR	08-NOV-2000	2000US-0246610
PR	08-NOV-2000	2000US-0246611
PR	08-NOV-2000	2000US-0246613
PR	17-NOV-2000	2000US-0249207
PR	17-NOV-2000	2000US-0249208
PR	17-NOV-2000	2000US-0249209
PR	17-NOV-2000	2000US-0249210
PR	17-NOV-2000	2000US-0249211
PR	17-NOV-2000	2000US-0249212
PR	17-NOV-2000	2000US-0249213
PR	17-NOV-2000	2000US-0249214
PR	17-NOV-2000	2000US-0249215
PR	17-NOV-2000	2000US-0249216
PR	17-NOV-2000	2000US-0249217
PR	17-NOV-2000	2000US-0249218
PR	17-NOV-2000	2000US-0249244
PR	17-NOV-2000	2000US-0249245
PR	17-NOV-2000	2000US-0249264
PR	17-NOV-2000	2000US-0249265
PR	17-NOV-2000	2000US-0249297
PR	17-NOV-2000	2000US-0249299
PR	17-NOV-2000	2000US-0250160
PR	01-DEC-2000	2000US-0250300
PR	01-DEC-2000	2000US-0250391
PR	05-DEC-2000	2000US-0251030
PR	05-DEC-2000	2000US-0251988
PR	05-DEC-2000	2000US-0256719
PR	06-DEC-2000	2000US-0251479
PR	08-DEC-2000	2000US-0251856
PR	08-DEC-2000	2000US-0251866
PR	08-DEC-2000	2000US-0251869
PR	08-DEC-2000	2000US-0251899
PR	11-DEC-2000	2000US-0251990
PR	05-JAN-2001	2001US-0259678

PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Rosen CA, Barash SC, Ruben SM;
XX	
DR	WPI; 2001-465566/50.
DR	N-PSDB; AAS41357.
XX	
PT	Novel polypeptides and polynucleotides useful for diagnosing,
PT	preventing, treating neural, immune system, muscular, reproductive,
PT	pulmonary, cardiovascular, renal, proliferative disorders and cancerous
XX	diseases -
XX	
PS	claim 11; SEQ ID NO 1483; 1180bp; English.
XX	
CC	The present invention relates to the isolation of novel human enzyme

polypeptides), and the cDNA (AA540735-AA541684) and genomic sequences encoding them. The enzyme polyphosphatases of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of a wide range of disorders including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. haemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. influenza). The polynucleotides of the invention can also be used in gene therapy. AA122915-AA23814 represent the novel human enzyme polypeptides of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pct/sequences](http://wipo.int/pub/published_pct/sequences).

SQ Sequence 298 AA;

Query Match	10.3%	Score 6	DB 22	Length 298
Best Local	Similarity 100.0%	Pred. No.	2e+02	
Matches 6	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY 26 FEDLYD 31

Db 184 FEDLYD 185

RESULT 40

ID	AAM25792	standard; Protein; 298 AA.
----	----------	----------------------------

AC AAM25792;

DT 16-OCT-2001 (first entry)

DE	Human protein sequence	SEQ ID NO:1307.

KM Human; ulcer; HIV infection; human immunodeficiency virus;
KM antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KM antibacterial; endocrine; cardiac; central nervous system; virucide;
KM anti-HIV; fungicide; antimutagen; cardiovascular; antileukemic; anaemia;
KM antiaggregant; haemostatic; vulnery; anticancer; osteopathic; eczema;
KM dermatological; antiallergic; antiaslatic; antidiabetic; cystostatic;
KM neuroprotective; antidepressant; nootropic; antiparkinsonian; infection
KM immunostimulant; gene therapy; autisense therapy; vaccine; inflammation
KM antidiaphoretic; rheumatoid arthritis; septic shock; pancreatitis;
KM cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KM genetic disease; haematopoietic disease; platelet disorder; asthma;
KM thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KM allergic rhinitis; diabetes; multiple sclerosis; depression;
KM Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KM neurological disorder.

OS Homo sapiens.

PN W0200153455-A2.

PD 26-JUL-2001.

PF 22-DEC-2000; 2000WO-US35017.

PR 23-DEC-1999; 99US-0471275.

PR 25-APR-2000; 2000US-0552317.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-457603/49.

DR N-PSDB; AAH99733.
XX Isolated human polynucleotides encoding polypeptides, useful for the
PT treatment and diagnosis of e.g. cancer, ulcers and HIV Infection -
XX
PS Claim 20; Page 271; 1217pp; English.
XX
XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC AAM25963. The proteins can have activities based on the tissues and
CC cells they are expressed in, such as: antiinflammatory; antirheumatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; virucide; anti-HIV; fungicide; antitumagen;
CC cardiovascular; antianaemic; antiaggregant; hemostatic; vulnerary;
CC antidiabetic; osteopathic; dermatological; antiallergic; antiasthmatic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders.
XX
XX
SO Sequence 298 AA;
Query Match 10.3%; Score 6; DB 22; Length 298;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 26 FEDLYD 31
Db 184 FEDLYD 189
RESULT 41
AAC08656
ID AAC08656 standard; Protein; 322 AA.
XX
AC AAC08656;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 6280.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137502.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145293.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159564.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 10.3%; Score 6; DB 21; Length 322;
Best Local Similarity 100.0%; Pred. No. 2,1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 KVVSY 38
|||||
Db 243 KVVSY 248

RESULT 42
AAU25613
ID AAU25613 standard; Protein; 336 AA.
XX
AC AAU25613;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human G Protein-Coupled Receptor (GPCR) polypeptide #60.
XX
KW Human; G-protein coupled receptor; GPCR; mental disorder; schizophrenia; attention deficit disorder; anxiety; depression; bipolar disorder; neurological disorder; Huntington's disease; dementia; obesity; anorexia; metabolic disorder; Parkinson's disease; Tourette's syndrome; thrombosis; type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer; cardiomyopathy; atherosclerosis; human immunodeficiency virus; HIV; viral infection; immunostimulant; neuroleptic; nootropic; tranquiliser; antidepressant; anorectic; gene therapy.
KW
XX
XX Homo sapiens.
OS
XX
PN W0200162797-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-US05676.
XX
PR 23-FEB-2000; 2000US-0184247.
PR 23-FEB-2000; 2000US-0184303.
PR 23-FEB-2000; 2000US-0184304.
PR 23-FEB-2000; 2000US-0184305.
PR 23-FEB-2000; 2000US-0184397.
PR 02-MAR-2000; 2000US-0186457.
PR 03-MAR-2000; 2000US-0186810.
PR 09-MAR-2000; 2000US-0188064.
PR 13-MAR-2000; 2000US-0188880.
PR 03-APR-2000; 2000US-0194344.
PR 23-JUN-2000; 2000US-0213861.
PR 11-JUL-2000; 2000US-0217369.
PR 11-JUL-2000; 2000US-0217370.

PR	14-JUL-2000; 2000US-0218337.
PR	20-JUL-2000; 2000US-0218492.
PA	(PHNA) PHARMACIA & UPJOHN CO.
PI	Vogel I G, Wood LS, Parodi LA, Lind P;
XX	
DR	WPI: 2001-570628/64.
DR	N-PSDB; AAS42865.
PS	Claim 35; Page 94; 279pp; English.
XX	
CC	Sequences AA025554-AA025616 represent human G-protein coupled receptor
CC	(GPCR) polypeptides of the invention. The proteins and their associated
CC	DNA sequences can be used to identify compounds which bind to GPCR
CC	polypeptides and in screening for compounds that modulate GPCR activity.
CC	by screening a human subject for the presence of mutations in GPCR DNA, a
CC	GPCR-related disorder or a genetic predisposition can be diagnosed. The
CC	sequences can also be used for treatment and prevention of mental
CC	disorders such as schizophrenia, attention deficit disorder, anxiety,
CC	depression, dementia and bipolar disorder, neurological disorders such as
CC	Huntington's disease, Parkinson's disease and Tourette's syndrome,
CC	metabolic disorders such as obesity, anorexia and type 2 diabetes,
CC	cardiovascular disorders such as thrombosis, myocardial infarction,
CC	cardiomyopathy and atherosclerosis, viral infections caused by HIV and
CC	cancers.
XX	
SO	Sequence 336 AA;
QY	41 YNLFTK 46
Db	321 YNLFTK 326
Query Match	10.3%; Score 6; DB 22; Length 336;
Best Local Similarity	100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
RESULT 43	
AAE08556	
ID	AAE08556 standard; Protein; 336 AA.
XX	
AC	AAE08556;
XX	
DT	15-NOV-2001 (first entry)
DE	
XX	Human seven transmembrane protein, 39406 protein.
XX	
Human; cytosolic; immunosuppressive; vasotropic; antiinflammatory;	
cardiac; haemostatic; vulnery; vitruide; hepatotropic; nephrotropic;	
neuroprotective; cerebroprotective; anticonvulsant; noctropic; oedema;	
leihsmaniasis; rheumatoid arthritis; systemic lupus erythematosus;	
splenoomegaly; Niemann-Pick disease; adult respiratory distress syndrome;	
asthma; diarrhoea; Crohn's disease; dysentery; jaundice; cholestasis;	
cirrhosis; Wilson's disease; glomerulonephritis; nephrotic syndrome;	
tumour; urinary tract infection; rhabdomyosarcoma; cerebral ischaemia;	
hypoxia; intracranial haemorrhage; acute meningitis; Parkinson's disease;	
Alzheimer's disease; glioma; stroke; Huntington's disease; osteoporosis;	
ricketts; osteonecrosis; Paget's disease; osteosarcoma; atherosclerosis;	
dwarfism; Kaposi sarcoma; angina pectoris; ischaemic heart disease;	
hypertension; myocardial infarction; hypertension; myocarditis; cancer;	
vaccine; gene therapy; 39406 protein; seven transmembrane protein; GPCR;	
G-protein-coupled receptor.	
XX	
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
FT	1..32
Peptide	/label=Signal_peptide
TT	

Protein	33..336	
Misc-difference	1	/note="This residue is given as Asn in the sequence shown as SEQ ID NO: 2 in figure 1 of the specification"
Domain	1..11	/label= Non-transmembrane_domain
Modified-site	3..6	/note="Asn is N-glycosylated"
Domain	12..35	/label= Transmembrane_domain
Modified-site	29..34	/label= N-myristoylation_site
Domain	36..51	/label= Non-transmembrane_domain
Modified-site	44..47	/label= Amidation_site
Modified-site	46..49	/note="cAMP and cGMP dependent protein kinase phosphorylation site"
Domain	52..76	/label= Transmembrane_domain
Domain	77..92	/label= Non-transmembrane_domain
Domain	93..115	/label= Transmembrane_domain
Domain	116..154	/label= Non-transmembrane_domain
Modified-site	126..129	/note="Casein Kinase II phosphorylation site"
Modified-site	132..135	/note="Casein Kinase II phosphorylation site"
Domain	155..177	/label= Transmembrane_domain
Modified-site	161..166	/label= N-myristoylation_site
Modified-site	175..178	/note="Casein Kinase II phosphorylation site"
Domain	178..193	/label= Non-transmembrane_domain
Modified-site	180..183	/note="Casein Kinase II phosphorylation site"
Domain	194..218	/label= Transmembrane_domain
Modified-site	195..200	/label= N-myristoylation_site
Modified-site	203..208	/label= N-myristoylation_site
Domain	219..250	/label= Non-transmembrane_domain
Modified-site	236..239	/note="Casein Kinase II phosphorylation site"
Modified-site	246..248	/note="Protein kinase C phosphorylation site"
Domain	251..268	/label= Transmembrane_domain
Modified-site	299..301	/note="Protein kinase C phosphorylation site"
Modified-site	302..305	/note="Casein Kinase II phosphorylation site"
Modified-site	315..317	/note="Protein kinase C phosphorylation site"
WO200159109-A1.		
16-AUG-2001.		
08-FEB-2001; 2001WO-US04074.		
08-FEB-2000; 2000US-0180912.		
(MILL-) MILLENNIUM PHARM INC.		

PI	Glucksmann MA, Galvin KM;
XX	
DR	WPI: 2001-522476/57.
XX	N-PSDB: AAD13260.
PT	Novel G-protein coupled receptor family polypeptide, 39406 polypeptide,
PT	useful as a target for diagnosis and treatment of 39406
PT	protein-mediated or -related disorders -
XX	
PS	Claim 8b; Fig 2; 125bp; English.
XX	
CC	The present sequence is 39406 protein, a seven transmembrane
CC	protein from human. 39406 protein is a receptor belonging to the
CC	superfamily of G-protein-coupled receptors (GPCR). The sequences of
CC	the invention of G-protein-coupled receptors for the diagnosis and treatment
CC	of 39406 protein-mediated or -related disorders, and for identifying
CC	agonists and antagonists for diagnosis and treatment. 39406 proteins
CC	are useful for treating disorders of spleen (e.g. splenomegaly, kala-
CC	azar, leishmaniasis, disorders associated with splenomegaly including
CC	infections, congestive states, lymphohaematogenous disorders, immunologic
CC	-inflammatory conditions such as rheumatoid arthritis and systemic lupus
CC	erythematosus, Gaucher's disease, mucopolysaccharidoses and Niemann-Pick
CC	disease), lung (e.g. congenital anomalies, pulmonary congestion, oedema,
CC	adult respiratory distress syndrome, haemorrhage, chronic obstructive
CC	pulmonary disease, goodpasture's syndrome, pulmonary hypertension and
CC	asthma), colon (e.g. enterocolitis such as diarrhoea and dysentery,
CC	viral gastroenteritis), bacterial enterocolitis, miscellaneous intestinal
CC	inflammatory disorders, drug-induced intestinal injury, idiopathic
CC	inflammatory bowel disease, Crohn's disease, tumours of colon and
CC	ulcerative colitis), liver (e.g. hepatic injury, jaundice, cholestasis,
CC	viral hepatitis, cirrhosis, Wilson's disease, autoimmune hepatitis and
CC	hepatic failure), kidney (e.g. glomerulonephritis, nephrotic syndrome,
CC	hereditary nephritis, urinary tract infection and acute tubular
CC	necrosis), skeletal muscle (e.g. tumours such as rhabdomyosarcoma),
CC	brain (e.g. hypoxia, cerebral ischaemia, intracranial haemorrhage,
CC	acute meningitis, Parkinson's disease, Alzheimer's disease, gliomas,
CC	chronic bacterial meningencephalitis, multiple sclerosis, amyotrophic
CC	lateral sclerosis, stroke and Huntington's disease), uterus and
CC	endometrium (e.g. inflammations, menopausal and post-menopausal changes
CC	and malignant tumours of endometrium), bones (osteoporosis, rickets,
CC	osteonecrosis, Paget's disease, osteosarcoma, type I collagen disease,
CC	dwarfism and metastatic tumours), blood vessels (e.g. atherosclerosis,
CC	vascular diseases, hypertension, tumours such as Kaposi sarcoma and
CC	disorders of veins and lymphatics), ovary and heart (e.g. heart failure,
CC	angina pectoris, ischaemic heart disease, myocardial infarction,
CC	hypertension, myocarditis and congenital heart disease) and cancers.
CC	39406 sequences are used as vaccines. They are also used in gene
CC	therapy.
XX	
XX	Sequence 336 AA;
SO	
Query Match	10.3%; Score 6; DB 22; Length 336;
Best Local Similarity	100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY 41 YNLFTR 46	
Db 321 YNLFTR 326	
RESULT 44	
ID AA004374	
XX AA004374 standard; Protein: 336 AA.	
XX AA004374;	
DT 23-OCT-2001 (first entry)	
DE Human G-protein coupled receptor, hRUP20.	
XX Human G-protein coupled receptor; GPCR; hRUP20; agonist;	
KM Inverse agonist; lung cancer.	
KM	
XX	

OS	Homo sapiens.
XN	WO200136471-A2.
PX	
PD	25-MAY-2001.
PF	16-NOV-2000; 2000MO-US31509.
XX	
PR	17-NOV-1999; 99US-0166088.
PR	17-NOV-1999; 99US-0166099.
PR	17-NOV-1999; 99US-0166369.
PR	23-DEC-1999; 99US-0171900.
PR	23-DEC-1999; 99US-0171901.
PR	23-DEC-1999; 99US-0171902.
PR	11-FEB-2000; 2000US-0181749.
PR	14-MAR-2000; 2000US-0189258.
PR	10-APR-2000; 2000US-0195898.
PR	10-APR-2000; 2000US-0195899.
PR	10-APR-2000; 2000US-0196078.
PR	28-APR-2000; 2000US-0200419.
PR	12-MAY-2000; 2000US-0203630.
PR	12-JUN-2000; 2000US-0210741.
PR	12-JUN-2000; 2000US-0210982.
PR	21-AUG-2000; 2000US-0226760.
PR	26-SEP-2000; 2000US-0235418.
PR	26-SEP-2000; 2000US-0235779.
PR	20-OCT-2000; 2000US-0242332.
PR	20-OCT-2000; 2000US-0242343.
XX	
PA	(AREN-) ARENA PHARM INC.
PI	
PI	Chen R, Dang HT, Lowitz KP;
XX	
DR	WPI: 2001-355616/37.
N-PSDB:	AAS07947.
PT	Endogenous and non-endogenous versions of human G-protein coupled receptors for direct identification of candidate compounds as agonists,
PT	Inverse agonists or partial agonists for use as therapeutic agents -
XX	
PS	Claim 49; Page 112-113; 160pp; English.
CC	The sequence represents a human G-protein coupled receptor (GPCR), hGPR20. The endogenous and non-endogenous, constitutively activated versions of human G-protein coupled receptors (GPCR), are useful for direct identification of candidate compounds as receptor agonists, CC inverse agonists or partial agonists having applicability as therapeutic agents for treating diseases related to GPCR, e.g. lung cancer. CC Non-endogenous version of human GPCRs are also utilized in research settings and in vitro and in vivo system, incorporating GPCRs can be CC utilised to elucidate and understand the roles these receptors play in the human condition, both normal and diseased.
SQ	Sequence 336 AA;
OY	Query Match 10.3%; Score 6; DB 22; Length 336; Best Local Similarity 100.0%; Pred.No.2.2e+02; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	41 YNLFTK 46 321 YNLFTK 326
RESULT 45	
ID	AA64299
AC	AA664299 standard; Protein; 336 AA.
XT	AA664299;
JT	21-SEP-2001 (first entry)

DE Human GTP-binding protein-coupled receptor GPRV77.
 XX
 XX GTP-binding protein-coupled receptor; neuroprotective; immunomodulatory;
 KM muscular; urinary; circulatory; anorectic; human; guanosine triphosphate;
 KM G-protein.
 XX
 OS Homo sapiens.
 XX
 XX WO200148189-A1.
 XX
 XX
 PD 05-JUL-2001.
 XX
 XX 28-DEC-2000; 2000MO-JP09409.
 PF
 XX 28-DEC-1999; 99JP-0375152.
 PR 31-MAR-2000; 2000JP-0101339.
 PR 23-MAY-2000; 2000JP-0155978.
 XX
 XX (HELI-) HELIX RES INST.
 PA
 XX
 XX
 PI Matsumoto S, Oda T, Saito Y, Morikawa N, Yoshida K, Suwa M;
 PI Sugiyama T;
 XX
 DR WPI: 2001-425663/45.
 DR N-PSDB; AAA49332.
 XX
 XX Family of guanosine triphosphate binding protein coupled receptors and
 PT genes encoding them for treatment and prevention of diseases associated
 PT with these receptors -
 PT
 XX
 PS Claim 1; Pages 112-114; 137pp; Japanese.
 CC
 CC The present sequence is the protein sequence for a human guanosine
 CC triphosphate (GTP)-binding protein-coupled receptor. The receptor is
 CC useful for the investigation, diagnosis, treatment and prevention of
 CC diseases associated with GTP-binding protein-coupled receptors, including
 CC neurological, circulatory, digestive system, immune system, muscle and
 CC urinary system disorders. GTP-binding proteins are also known as
 CC G-proteins.
 CC
 XX
 SQ Sequence 336 AA;

Query Match 10.3%; Score 6; DB 22; Length 336;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 41 YNLFTR 46
 |||||
 Db 321 YNLFTR 326

Search completed: November 5, 2002, 05:24:41
 Job time : 68 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 05:05:29 ; Search time 44 Seconds
(Without alignments)
126.663 Million cell updates/sec

Title: US-09-727-892a-99

Perfect score: 58

Sequence: 1 MERKRYTVLLYCDEIKGHFP.....YEYNLFTKKYATIEYKEI 58

Scoring table: OLIGO

Searched: 283138 seqs, 96089334 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

PIR_71:*
1: piri:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	12.1	348	2	A32512 glycerol-3-phospha
2	7	12.1	433	2	S26646 transcription fact
3	6	10.3	106	2	H84562 hypothetical prote
4	6	10.3	127	2	H90460 conserved hypotet
5	6	10.3	162	2	H96705 hypothetical prote
6	6	10.3	170	2	AF1379 hypothetical prote
7	6	10.3	170	2	AG1748 hypothetical prote
8	6	10.3	173	2	F91252 probable tail fibe
9	6	10.3	178	2	G86788 hypothetical prote
10	6	10.3	186	2	H75004 hypothetical prote
11	6	10.3	186	2	A71158 hypothetical prote
12	6	10.3	186	2	T21961 hypothetical prote
13	6	10.3	199	2	T27239 hypothetical prote
14	6	10.3	201	2	G83890 hypothetical prote
15	6	10.3	205	2	D69870 conserved hypotet
16	6	10.3	216	2	S48486 probable membra
17	6	10.3	232	2	A87504 6-phospho-glucono-
18	6	10.3	235	2	B71613 RAB GTPase PRB0500
19	6	10.3	241	2	B69655 two-component resp
20	6	10.3	247	2	C87423 cytochrome c oxida
21	6	10.3	253	2	H70380 conserved hypotet
22	6	10.3	255	2	F86203 hypothetical prote
23	6	10.3	259	2	G81427 periplasmic protei
24	6	10.3	264	2	T37246 probable transcrip
25	6	10.3	268	2	B72352 conserved hypotet
26	6	10.3	273	2	S20069 ribonucleoprotein
27	6	10.3	274	2	D95339 hypothetical prote
28	6	10.3	274	2	G72685 probable alanyl-tr
29	6	10.3	279	2	S26203 RNA-binding protei

30	6	10.3	288	2	T44603 hypothetical prote
31	6	10.3	291	2	S20070 ribonucleoprotein
32	6	10.3	292	2	S26204 RNA-binding protei
33	6	10.3	293	2	D81896 Neisseria meningit
34	6	10.3	294	2	T05725 op31MHV protein
35	6	10.3	301	2	G97187 UDP-glucose 4-epim
36	6	10.3	307	2	A71057 probable sulfatase
37	6	10.3	320	2	AD0563 ferrochelatase (im
38	6	10.3	327	2	G90139 deacetylase, proba
39	6	10.3	336	2	T19757 hypothetical prote
40	6	10.3	339	2	T49597 hypothetical prote
41	6	10.3	346	2	C82435 conserved hypotet
42	6	10.3	361	2	T49881 pectin methyl-este
43	6	10.3	377	2	G97314 alcohol dehydrogen
44	6	10.3	378	2	S71201 biotin synthase (E
45	6	10.3	388	2	H64427 hypothetical prote

ALIGNMENTS

RESULT 1
A32512 glycerol-3-phosphate dehydrogenase (NAD+) (EC 1.1.1.8) - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 04-Sep-1998
C:Accession: A32512
R:Otto, J.; Machleidt, W.; Wächter, E.; Rueckl, G.; Machleidt, I.
unpublished results, cited by Otto, J., Argos, P., and Rossmann, M.G. Eur. J. Biochem
A:Reference number: A32512
A:Accession: A32512
A:Molecule type: Protein
A:Residues: 1-348 <ONT>
C:Superfamily: glycerol-3-phosphate dehydrogenase
C:Keywords: NAD; oxidoreductase

Query Match 12.1%; Score 7; DB 2; Length 348;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 CDEIKGH 18
DB 101 CDEIKGH 107

RESULT 2
S26646 transcription factor IIE - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
C:Accession: S26646
R:Ohkuma, Y.; Hashimoto, S.; Roeder, R.G.; Horikoshi, M.
Nucleic Acids Res. 20, 5838, 1992
A:Title: Identification of two large subdomains in TFIIIE-alpha on the basis of homolo
A:Reference number: S26646; MUID:93087200
A:Accession: S26646
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-433 <OHK>
A:Cross-references: EMBL:Z14131; NID:965130; PIDN:CA78505.1; PID:965131

Query Match 12.1%; Score 7; DB 2; Length 433;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 MFEDLYD 31
DB 427 MFEDLYD 433

RESULT 3
H84562 hypothetical protein At2g18320 [Imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: H84562
A:Status: preliminary
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: H84562
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-106 <STO>
A:Cross-references: GB:AE002093; NID:94309735; PIDN:AD15505.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g18320
A:Map position: 2

Query Match 10.3%; Score 6; DB 2; Length 106;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 KYAYII 52
|||||
DB 71 KYAYII 76

RESULT 4
H90460
conserved hypothetical protein [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C:Accession: H90460
A:Status: preliminary
A:Title: R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyaz, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
submitted to Genbank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: H90460
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-127 <NR>
A:Cross-references: GB:AE006641; NID:913816188; PIDN:AAK42943.1; GSPDB:GN00155
C:Genetics:
A:Gene: SS02833
C:Superfamily: Aeropyrum pernix hypothetical protein APE1561

Query Match 10.3%; Score 6; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 KVVYSY 38
|||||
DB 32 KVVYSY 37

RESULT 5
H96705
hypothetical protein T22E19.13 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: H96705
A:Status: preliminary
A:Title: Logg, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719
A:Accession: H96705
A:Status: preliminary
A:Title: Comparative genomics of Listeria species.
A:Residues: 1-162 <STO>
A:Cross-references: GB:AE005173; NID:96715728; PIDN:AAF26489.1; GSPDB:GN00141
C:Genetics:
A:Gene: T22E19.13
A:Map position: 1

Query Match 10.3%; Score 6; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 IIEYIK 56
|||||
DB 87 IIEYIK 92

RESULT 6
AF1379
hypothetical protein lmo2438 [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AF1379
A:Status: preliminary
A:Title: Jones, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
r; Glaser, P.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.;
Ok, C.; Schluter, T.; Simoes, N.; Tlertez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AF1379
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-170 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAD00516.1; PID:916411926; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2438

Query Match 10.3%; Score 6; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 LFTKKY 48
|||||
DB 127 LFTKKY 132

RESULT 7
AG1748
hypothetical protein lln2532 [imported] - Listeria innocua (strain C11p11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AG1748
A:Status: preliminary
A:Title: Jones, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
r; Glaser, P.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.;
Ok, C.; Schluter, T.; Simoes, N.; Tlertez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1748
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-170 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC97759.1; PID:916415054; GSPDB:GN00178
A:Experimental source: strain C11p11262
C:Genetics:
A:Gene: lln2532

Query Match 10.3%; Score 6; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 LFTKKY 48
|||||
DB 127 LFTKKY 132

RESULT 8

F91252
Probable tail fiber assembly protein [imported] - Escherichia coli (strain O157:H7, subs
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: F91252
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhnara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: F91252
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-173 <HAV>
A:Cross-references: GB:BA000007; PIDN:BA038413.1; PID:913364466; GSPDB:GN00154
C:Experimental source: strain O157:H7, substrain RMD 0509552
C:Genetics:
A:Gene: EC84990

Query Match 10.3%; Score 6; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 LFTKKY 48
|||||
DB 37 LFTKKY 42

RESULT 9

G86788
Hypothetical protein ynef [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: G86788
R:Boletín, A.; Winkler, P.; Mauger, S.; Jallón, O.; Malarne, K.; Weissenbach, J.; Ehrlich
genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: G86788
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-178 <STO>
A:Cross-references: GB:AE005176; PID:912724290; PIDN:AAK05409.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: ynef

Query Match 10.3%; Score 6; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 LFTKKY 48
|||||
DB 136 LFTKKY 141

RESULT 10

H75004
Hypothetical protein PAB1318 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: H75004

R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome s
A:Reference number: A75001
A:Accession: H75004
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-186 <KAW>
A:Cross-references: GB:AJ248288; GB:AL096836; NID:95458960; PIDN:CAB50478.1; PID:9545
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB1318
C:Superfamily: conserved hypothetical protein M0541

Query Match 10.3%; Score 6; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 53 EYKEI 58
|||||
DB 152 EYKEI 157

RESULT 11

A71158
Hypothetical protein PH0464 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000
C:Accession: A71158
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Se
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Ogu
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophil
A:Reference number: A71000; MUID:98344137
A:Accession: A71158
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-186 <KAW>
A:Cross-references: NID:93236129; PIDN:BAA29550.1; PID:93256867
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenB
C:Genetics:
A:Gene: PH0464
C:Superfamily: conserved hypothetical protein M0541

Query Match 10.3%; Score 6; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 53 EYKEI 58
|||||
DB 152 EYKEI 157

RESULT 12

T21961
Hypothetical protein F38C2.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21961
R:Mortimore, B.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19494
A:Accession: T21961
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-186 <WIL>
A:Cross-references: EMBL:Z82267; PIDN:CAB05191.1; GSPDB:GN00022; CESP:F38C2.5
A:Experimental source: clone F38C2
C:Genetics:
A:Gene: CESP:F38C2.5
A:Map position: 4
A:Introns: 107/3

Query Match 10.3%; Score 6; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KYKTVL 9
|||||
DB 115 KYKTVL 120

RESULT 13

T27239
hypothetical protein Y57G11C.25 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27239

R:McMurray, A.
submitted to the EMBL Data Library, September 1997

A:Reference number: Z20330

A:Accession: T27239

A:Status: preliminary; translated from GB/EMBL/DDAJ

A:Molecule type: DNA

A:Residues: 1-199 <NTL>

A:Cross-references: EMBL:Z299281; PIDN:CAB16528.1; GSPDB:GN00022; CESP:Y57G11C.25

A:Experimental source: clone Y57G11C

C:Genetics:

A:Gene: CESP:Y57G11C.25

A:Map position: 4

A:Introns: 103/3

Query Match 10.3%; Score 6; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KYKTVL 9
|||||
DB 111 KYKTVL 116

RESULT 14

G83890
hypothetical protein BH1927 [imported] - *Bacillus halodurans* (strain C-125)

C:Species: *Bacillus halodurans*

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2001

C:Accession: G83890

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: G83890

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-201 <STO>

A:Cross-references: GB:AP001513; GB:BA000004; NID:910174345; PIDN:BA05646.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH1927

C:Superfamily: *Bacillus subtilis* conserved hypothetical protein yest

Query Match 10.3%; Score 6; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 FTKRYA 49
|||||
DB 195 FTKRYA 200

RESULT 15

D69870

conserved hypothetical protein ykYA - *Bacillus subtilis*

N:Alternate names: hypothetical protein (acea 5' region)

C:Species: *Bacillus subtilis*

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999

C:Accession: D69870; A36718

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber

C: Brod, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;

A: Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal

lechi, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M

koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanl

A:Authors: Schleich, S.; Schroeter, R.; Scofield, F.; Sekiguchi, J.; Sekowska, A.; Se

akenuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*

A:Reference number: A69580; MUID:98044033

A:Accession: D69870

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-205 <KUN>

A:Cross-references: GB:Z99111; GB:AL009126; NID:q2633699; PIDN:CAB13330.1; PID:ell1850

A:Experimental source: strain 168

A:Hemilae, H.; Palva, A.; Paulin, L.; Arvidson, S.; Palva, I.

J. Bacteriol. 172, 5052-5063, 1990

A:Title: Secretory S complex of *Bacillus subtilis*: sequence analysis and identity to

A:Reference number: A36718; MUID:90368558

A:Accession: A36718

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 186-205 <HEM>

A:Cross-references: GB:M31542

C:Genetics:

A:Gene: ykYA

Query Match 10.3%; Score 6; DB 2; Length 205;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 EDLYDA 32
|||||
DB 192 EDLYDA 197

RESULT 16

S48486
probable membrane protein YIR024c - yeast (*Saccharomyces cerevisiae*)

C:Species: *Saccharomyces cerevisiae*

C>Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 05-Nov-1999

C:Accession: S48486

R:Rowley, K.
submitted to the EMBL Data Library, October 1994

A:Reference number: S48478

A:Accession: S48486

A:Molecule type: DNA

A:Residues: 1-216 <ROW>

A:Cross-references: GB:Z47047; EMBL:Z38061; NID:9603997; PID:q763369; GSPDB:GN00009;

C:Genetics:

A:Gene: MIPS:YIR024c

A:Map position: 98

C:Keywords: transmembrane protein

E:47-63/Domain: transmembrane #status predicted <TMM>

Query Match 10.3%; Score 6; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MERKYK 6
|||||
DB 71 MERKYK 76

RESULT 17
A87504
6-phospho-glucono-lactonase [imported] - *Caulobacter crescentus*
C:Species: *Caulobacter crescentus*
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: A87504
R:Nierman, W.C.; Felblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; Debey, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolotn, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: A87504
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-232 <STO>
A:Cross-references: GB:AE005673; NID:913423535; PIDN:AAK24029.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC2056

Query Match 10.3%; Score 6; DB 2; Length 232;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 27 EDLYDA 32
|||||
Db 15 EDLYDA 20

RESULT 18
B71613
RAB GPase PRB0500C - malaria parasite (*Plasmodium falciparum*)
C:Species: *Plasmodium falciparum*
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 02-Feb-2001
C:Accession: B71613
R:Gardner, M.J.; Tetteh, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.; Perlita, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum*.
A:Reference number: A71600; MUID:99021743
A:Accession: B71613
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-235 <GAR>
A:Cross-references: GB:AE001399; GB:AE001362; NID:93845200; PIDN:AACT1889.1; PID:9384520
A:Experimental source: clone 307
C:Genetics:
A:Gene: PRB0500C
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: nucleotide binding; P-loop
F:13-20/Region: nucleotide-binding motif A (P-loop)
F:154-157/Region: GTP-binding NKXD motif
F:183-185/Region: GTP-binding SAK/L motif

Query Match 10.3%; Score 6; DB 2; Length 235;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 KYIVL 10
|||||
Db 7 KYIVL 12

RESULT 19
B69655
Two-component response regulator *lytR*-involved - *Bacillus subtilis*
C:Species: *Bacillus subtilis*
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: B69655
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berta, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chn, A.; Ehlich, S.D.; Emmerison, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galiech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holstappel, S.; Hosono, S.; Hullo, M. Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, R.; Masuda, S.; Mau Y, M.; Ogawa, K.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se akenchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya T.; Winers, P.; Wipat, A.; Yamamoto, H.; Yaman, K.; Yasumoto, K.; Yata, K.; Yoshida A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*
A:Reference number: A69580; MUID:98044033
A:Accession: B69655
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-241 <KUN>
A:Cross-references: GB:299118; GB:AL009126; NID:92635200; PIDN:CA14852.1; PID:926353
A:Experimental source: strain 168
C:Genetics:
A:Gene: *lytR*
C:Superfamily: *yejH* protein; response regulator homology
C:Keywords: phosphoprotein
F:4-113/Domain: response regulator homology <RRH>
F:54/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 10.3%; Score 6; DB 2; Length 241;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 53 EYKEI 58
|||||
Db 202 EYKEI 207

RESULT 20
C87423
cytochrome c oxidase, CcoO subunit [imported] - *Caulobacter crescentus*
C:Species: *Caulobacter crescentus*
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: C87423
R:Nierman, W.C.; Felblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; Debey, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: C87423
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-247 <STO>
A:Cross-references: GB:AE005673; NID:913422761; PIDN:AAK23383.1; GSPDB:GN00148
C:Genetics:
A:Gene: CCI402
C:Superfamily: Rhizobium cytochrome-c oxidase fixO chain

Query Match 10.3%; Score 6; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 YDAKV 35
|||||
Db 200 YDAKV 205

RESULT 21
H70380
conserved hypothetical protein *ag_933* - *Aquifex aeolicus*
C:Species: *Aquifex aeolicus*
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: H70380
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; V.
Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666
A:Accession: H70380
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-253 <NOF>
A:Cross-references: GB:AE000714; NID:g2983446; PIDN:AAC07040.1; PID:g2983459; GB:AE00065
A:Experimental source: strain VF5
C:Gene: aq_933
C:Superfamily: conserved hypothetical protein y100; conserved hypothetical protein y100

Query Match 10.3%; Score 6; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 EYKEI 58
|||||
DB 82 EYKEI 87

RESULT 22
F86203
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F86203
R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso, C.H.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: F86203
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-255 <STO>
A:Cross-references: GB:AE005172; NID:g7523696; PIDN:AAF63135.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 10.3%; Score 6; DB 2; Length 255;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 VVYSY 39
|||||
DB 189 VVYSY 194

RESULT 23
G81427
periplasmic protein Cj0111 [imported] - Campylobacter jejuni (strain NCTC 11168)
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Mar-2000
C:Accession: G81427
R:Parhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barré-Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypA:Reference number: A81250; MUID:20150912
A:Accession: G81427
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-259 <PAR>
A:Cross-references: GB:AL139074; GB:AL111168; NID:g6967505; PIDN:CAB72595.1; PID:g696760
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:

A:Gene: Cj0111
Query Match 10.3%; Score 6; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 LYDAKV 34
|||||
DB 218 LYDAKV 223

RESULT 24
T37246
probable transcription factor pos-1 - Caenorhabditis elegans
N:Alternate names: cytoplasmic zinc-finger protein
C:Species: Caenorhabditis elegans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C:Accession: T37246
R:Tabata, H.; Hill, R.J.; Mello, C.C.; Pless, J.R.; Kohara, Y.
Development 126, 1-11, 1999
A:Title: Pos-1 encodes a cytoplasmic zinc-finger protein essential for germline spectA:Reference number: Z21643; MUID:99054964
A:Accession: T37246
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-264 <TAB>
A:Cross-references: EMBL:AB006208; NID:g3767589; PIDN:BAA33854.1; PID:g3767590
C:Genetics:
A:Gene: pos-1
A:Map position: V
A:Function:
A:Description: required for germline specification by regulating expression of maternC:Keywords: DNA binding; transcription factor; transcription regulation

Query Match 10.3%; Score 6; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KYKTVL 9
|||||
DB 141 KYKTVL 146

RESULT 25
B72352
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: B72352
R:Nelson, K.E.; Clayton, R.A.; Gall, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; HicGarrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genomeA:Reference number: A72200; MUID:99287316
A:Accession: B72352
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-268 <ARN>
A:Cross-references: GB:AE001738; GB:AE000512; NID:g4981158; PIDN:AAD35735.1; PID:g498A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0651
C:Superfamily: Methanobacterium thermoautotrophicum conserved hypothetical protein MT

Query Match 10.3%; Score 6; DB 2; Length 268;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 IIEYIK 56
|||||
DB 91 IIEYIK 96

RESULT 26
S20069
ribonucleoprotein A, 29k - wood tobacco
C:Species: Nicotiana sylvestris (wood tobacco)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Sep-1999
C:Accession: S20069
R:Ye, L.; Li, Y.; Fukami-Kobayashi, K.; Go, M.; Konishi, T.; Watanabe, A.; Sugiyura, M.
Nucleic Acids Res. 19, 6485-6490, 1991
A:Title: Diversity of a ribonucleoprotein family in tobacco chloroplasts: two new chloro
A:Reference number: S20069; MUID:92093607
A:Accession: S20069
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-273 <YEL>
A:Cross-references: EMBL:X61113; NID:919753; PIDN:CAA3427.1; PID:919754
C:Superfamily: unassigned ribonucleoprotein repeat containing proteins; ribonucleoprotei
F:88-155/Domain: ribonucleoprotein repeat homology <RMI>
F:189-256/Domain: ribonucleoprotein repeat homology <RMI>
Query Match 10.3%; Score 6; DB 2; Length 273;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 31 DAKVY 36
|||||
Db 216 DAKVY 221
RESULT 27
D95339
hypothetical protein Sma1147 [imported] - Sinorhizobium meliloti (strain 1021) magaplast
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: D95339
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: D95339
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-274 <KUR>
A:Cross-references: GB:AE006469; PIDN:AAK65278.1; PID:g14523731; GSPDB:GN00165
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Voitholler, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Sma1147
A:Genome: plasmid
Query Match 10.3%; Score 6; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 YKTVLL 10
|||||
Db 3 YKTVLL 8
RESULT 28
G72685
probable alanyl-tRNA synthetase APE0903 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: G72685
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339
A:Accession: G72685
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-274 <KAM>
A:Cross-references: DDBJ:AP000060; NID:95104188; PIDN:BA079887.1; PID:d1043673; PID:g
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0903
Query Match 10.3%; Score 6; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 KTVLLY 11
|||||
Db 27 KTVLLY 32
RESULT 29
S26203
RNA-binding protein 30 - curled-leaved tobacco
C:Species: Nicotiana plumbaginifolia (curled-leaved tobacco)
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 24-Sep-1999
C:Accession: S26203; S19856
R:Misszack, M.; Klahre, U.; Levy, J.H.; Goodall, G.J.; Filipowicz, W.
Mol. Gen. Genet. 234, 390-400, 1992
A:Title: Multiple plant RNA binding proteins identified by PCR: expression of cDNAs e
A:Reference number: S26203; MUID:93024312
A:Accession: S26203
A:Molecule type: mRNA
A:Residues: 1-279 <MIE>
A:Cross-references: EMBL:X65118; NID:919707; PIDN:CAA46234.1; PID:g19708
C:Superfamily: unassigned ribonucleoprotein repeat containing proteins; ribonucleopro
F:88-155/Domain: ribonucleoprotein repeat homology <RMI>
F:195-262/Domain: ribonucleoprotein repeat homology <RMI>
Query Match 10.3%; Score 6; DB 2; Length 279;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 31 DAKVY 36
|||||
Db 222 DAKVY 227
RESULT 30
T44603
hypothetical protein CGI-83 [imported] - human
C:Species: Homo sapiens (man)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C:Accession: T44603
R:Lin, W.C.
submitted to the EMBL Data Library, May 1999
A:Description: Comparative gene cloning: Identification of novel human genes with Cae
A:Reference number: Z22808
A:Accession: T44603
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-288 <LIN>
A:Cross-references: EMBL:AF151841; PIDN:AAD34078.1
C:Genetics:
A:Map position: 8
Query Match 10.3%; Score 6; DB 2; Length 288;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 26 FEDLYD 31
|||||
DB 174 FEDLYD 179

RESULT 31

S20070
ribonucleoprotein B, 29K - wood tobacco
C:Species: Nicotiana sylvestris (wood tobacco)
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Sep-1999
C:Accession: S20070
R:Ye, L.; Li, Y.; Fukami-Kobayashi, K.; Go, M.; Konishi, T.; Watanabe, A.; Sugitara, M.
Nucleic Acids Res. 19, 6485-6490, 1991
A:Title: Diversity of a ribonucleoprotein family in tobacco chloroplasts: two new chloro
A:Reference number: S20069; MUID:92093607
A:Accession: S20070
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-291 <YEL>
A:Cross-references: EMBL:X61114; NID:g14134; PID:CAA43428.1; PID:g14135
C:Genetics:
A:Introns: 117/3; 151/3; 271/3
C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein
F:88-155/Domain: ribonucleoprotein repeat homology <RRM1>
F:208-275/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 10.3%; Score 6; DB 2; Length 291;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 DAKVY 36
|||||
DB 235 DAKVY 240

RESULT 32

S26204
RNA-binding protein 31 - curled-leaved tobacco
C:Species: Nicotiana glauca (curled-leaved tobacco)
C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 24-Sep-1999
C:Accession: S26204; S19857
R:Mieszcak, M.; Klahre, U.; Levy, J.H.; Goodall, G.J.; Filipowicz, W.
Mol. Gen. Genet. 234, 390-400, 1992
A:Title: Multiple plant RNA binding proteins identified by PCR: expression of cDNAs encod
A:Reference number: S26203; MUID:93024312
A:Accession: S26204
A:Molecule type: mRNA
A:Residues: 1-292 <MIE>
A:Cross-references: EMBL:X65117; NID:g19709; PID:CAA46233.1; PID:g19710
C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein
F:89-156/Domain: ribonucleoprotein repeat homology <RRM1>
F:209-276/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 10.3%; Score 6; DB 2; Length 292;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 DAKVY 36
|||||
DB 236 DAKVY 241

RESULT 33

D81896
Neisseria meningitidis hypothetical protein NMA1282 - Neisseria meningitidis (strain Z24
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: D81896
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churche, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: AB1775; MUID:20222556

A:Accession: D81896
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-293 <PAR>
A:Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PID:CA84534.1; PID:g737
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA1282
C:Superfamily: Neisseria meningitidis hypothetical protein NMA1282

Query Match 10.3%; Score 6; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 46 KKVAYI 51
|||||
DB 240 KKVAYI 245

RESULT 34

T05725
GPI1AHV protein - barley
C:Species: Hordeum vulgare (barley)
C>Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 21-Jan-2000
C:Accession: T05725
R:Churin, Y.; Hess, W.; Boerner, T.
Submitted to the EMBL Data Library, June 1998
A:Description: Characterization of three differently expressed cDNAs encoding chlorop
A:Reference number: Z15435
A:Accession: T05725
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-294 <CHU>
A:Cross-references: EMBL:AJ005286; NID:el318689; PID:CAA06469.1; PID:el318690
C:Experimental source: cv. Haisa, leaf
C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleopro
F:117-184/Domain: ribonucleoprotein repeat homology <RRM>

Query Match 10.3%; Score 6; DB 2; Length 294;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 DAKVY 36
|||||
DB 238 DAKVY 243

RESULT 35

G97187
UDP-glucose 4-epimerase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: G97187
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: G97187
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-301 <KUR>
A:Cross-references: GB:AE001437; PID:AAK80290.1; PID:g15025343; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC2334
C:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homolo

Query Match 10.3%; Score 6; DB 2; Length 301;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29 LYDAKV 34

Db 86 LYDAKV 91

RESULT 36

A:1057
probable sulfatase - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000

C:Accession: A71057

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin

M.; Ohtuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi

DNA Res. 5, 55-76, 1998

A>Title: Complete sequence and gene organization of the genome of a hyper-thermophilic

A:Reference number: A71000; MUID:98344137

A:Accession: A71057

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-307 <KAM>

A:Cross-references: GB:AP000005; NID:g3236132; PIDN:BAA30251.1; PID:g3257568

A:Experimental source: strain OT3

A>Note: this accession replaces an interim accession for a sequence replaced by GenBank

C:Genetics:

C:Superfamily: conserved hypothetical protein MJ1502

Query Match 10.3%; Score 6; DB 2; Length 307;

Best Local Similarity 100.0%; Pred. No. 85;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 AKVYVS 37

Db 201 AKVYVS 206

RESULT 37

AD0563

ferrochelatase [imported] - Salmonella enterica subsp. enterica serovar Typhl (strain CT

C:Species: Salmonella enterica subsp. enterica serovar Typhl

A>Note: this species has also been called Salmonella typhl

C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001

C:Accession: AD0563

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connelton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

S.; Moule, S.; O'Garra, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; PMID:11677608

A:Accession: AD0563

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-320 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD04974.1; PID:g16501760; GSPDB:GN00176

C:Genetics:

A:Gene: STY0533

Query Match 10.3%; Score 6; DB 2; Length 320;

Best Local Similarity 100.0%; Pred. No. 88;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 KKYAYI 51

Db 294 KKYAYI 299

RESULT 38

G90139

deacetylase, probable [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001

C:Accession: G90139

R:She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: G90139

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-327 <KUR>

A:Cross-references: GB:AE006641; NID:g13813134; PIDN:AAK40374.1; GSPDB:GN00155

C:Genetics:

A:Gene: SS00009

C:Superfamily: acetylpolymine aminohydrolyase; RPD3/acuc homology

Query Match 10.3%; Score 6; DB 2; Length 327;

Best Local Similarity 100.0%; Pred. No. 90;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 FTKRYA 49

Db 264 FTKRYA 269

RESULT 39

T19757

hypothetical protein C35D6.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T19757; T19758

R:Mathews, L.

submitted to the EMBL data library, November 1996

A:Reference number: Z19174

A:Accession: T19757

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-336 <WIL>

A:Cross-references: EMBL:Z82261; PIDN:CAB05144.1; GSPDB:GN00022; CESP:C35D6.1

A:Experimental source: clone C35D6

A:Accession: T19758

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-336 <WIL>

A:Cross-references: EMBL:Z82261; PIDN:CAB05145.1; GSPDB:GN00022; CESP:C35D6.2

A:Experimental source: clone C35D6

C:Genetics:

A:Gene: CESP:C35D6.1; CESP:C35D6.2

A:Map position: 4

A:Insertions: 109/1; 171/3

C:Superfamily: Caenorhabditis hypothetical protein C49G7.2

Query Match 10.3%; Score 6; DB 2; Length 336;

Best Local Similarity 100.0%; Pred. No. 92;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 KYAYII 52

Db 98 KYAYII 103

RESULT 40

T49597

hypothetical protein B3E4.100 [imported] - Neurospora crassa

C:Species: Neurospora crassa

C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000

C:Accession: T49597

R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu

submitted to the Protein Sequence Database, May 2000

A:Reference number: Z25022

A:Accession: T49597

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-339 <SCH>

A:Cross-references: EMBL:AL355931; GSPDB:GN00116; NCSP:B3E4.100

A:Experimental source: BAC clone B3E4; strain OR74A

C:Genetics:

A:Gene: NCSP:B3E4.100

A:Map position: 6

A:Introns: 85/1; 112/2; 132/1; 176/1; 233/2; 252/3; 268/3

C:Superfamily: Neurospora crassa hypothetical protein B3E4.100

Query Match

10.3%; Score 6; DB 2; Length 339;

Best Local Similarity 100.0%; Pred. No. 93;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 HPPHOI 23

Db 106 HPPHOI 111

RESULT 41

C82435 conserved hypothetical protein VCA0629 [Imported] - Vibrio cholerae (strain N16961 serog

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: C82435

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;

charadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Baas, S.; Qian, H.; Dragol, I.; Sellers, R.

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: C82435

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-346 <HEI>

A:Cross-references: GB:AE004593; GB:AE003853; NID:g9658042; PIDN:AAF96530.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VCA0629

A:Map position: 2

Query Match

10.3%; Score 6; DB 2; Length 346;

Best Local Similarity 100.0%; Pred. No. 94;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 AKVYS 37

Db 141 AKVYS 146

RESULT 42

T49881 pectin methyl-esterase-like protein - Arabidopsis thaliana

N:Alternate names: protein T211.130

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000

C:Accession: T49881

R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bannroft, I.; Mewes, H.W.; Rudd, S.; Le

submitted to the Protein Sequence Database, April 2000

A:Reference number: Z24493.

A:Accession: T49881

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-361 <BEV>

A:Cross-references: EMBL:AL163912; GSPDB:GN00063; ATSP:T211.130

A:Experimental source: cultivar Columbia; BAC clone T211

C:Genetics:

A:Gene: ATSP:T211.130

A:Map position: 5

A:Introns: 68/3; 161/3; 229/3; 309/2

Query Match

10.3%; Score 6; DB 2; Length 361;

Best Local Similarity 100.0%; Pred. No. 98;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 KVVYS 38

|||||
Db 282 KVVYS 287

RESULT 43

397314 alcohol dehydrogenase [Imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001

C:Accession: G97314

R:Molling, J.; Breton, G.; Omelchenko, M.V.; Markova, K.S.; Zeng, Q.; Gibson, R.; L

3; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

7 Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: G97314

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-377 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK81306.1; PID:g15026459; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC3375

C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology

Query Match

10.3%; Score 6; DB 2; Length 377;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 53 EYKEI 58

Db 237 EYKEI 242

RESULT 44

571201

371201 biotin synthase (EC 2.8.1.6) bioB [similarity] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 16-Feb-2001

C:Accession: S71201; B84865

R>Weaver, L.M.; Nikolau, B.J.

submitted to the EMBL Data Library, April 1995

A:Description: Characterization of the Arabidopsis biotin synthase gene.

A:Reference number: S71201

A:Accession: S71201

A:Molecule type: DNA

A:Residues: 1-378 <WEA>

A:Cross-references: EMBL:U24147; NID:g1045315; PIDN:AAAB0226.1; PID:g1045316

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y

us, D.; Nierman, W.C.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon,

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487

A:Accession: B84865

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-378 <STO>

A:Cross-references: GB:AE002093; NID:g2288983; PIDN:AAAB4312.1; GSPDB:GN00139

C:Genetics:

A:Gene: bioB; AT2g43360

A:Map position: 2

A:Introns: 68/3; 128/3; 164/2; 233/1; 279/3

A:Function:

A:Description: catalyzes conversion of dehydrobiotin to biotin

A:Pathway: biotin biosynthesis

A>Note: last step

C:Superfamily: biotin synthetase

C:Keywords: 2Fe-2S; biotin biosynthesis; iron-sulfur protein; metalloprotein; sulfur

94,98,101,231/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match

10.3%; Score 6; DB 2; Length 378;

Best Local Similarity 100.0%; Pred. No. 1e+02;

	Matches	6;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	53	EXIKKI	58							
DB	158	EXIKKI	163							

RESULT 45

H64427

hypothetical protein MJ1025 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C:Accession: H64427

R:Butt, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
reson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A:Reference number: A64300; MUID:96337999

A:Accession: H64427

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-388 <BD>

A:Cross-references: GB:U67545; GB:L77117; NID:g1591680; PIDN:AA99029.1; PID:g1591681; T

C:Genetics:

A:Map position: REV956916-955750

A:Start codon: GTG

C:Superfamily: hypothetical protein MJ1025

Query Match	10.3%;	Score 6;	DB 2;	Length 388;
Best Local Similarity	100.0%;	Pred. No. 1e+02;		
Matches	6;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

QY	13	DEIKGH	18
DB	247	DEIKGH	252

Search completed: November 5, 2002, 05:26:32
Job time : 49 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 05:23:44 ; Search time 69 Seconds
(without alignments)
145.416 Million cell updates/sec

Title: US-09-727-892a-99

Perfect score: 58
Sequence: 1 MERKRYTLVLCDEIKHFP.....KEYNLFTRKYAVIIEYKEI 58

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size: 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp-phage:*
10: sp_plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	12.1	433	13	091859 xenopus lae
2	7	12.1	438	2	032519 eubacterium
3	7	12.1	677	5	09VE10 drosophila
4	7	12.1	844	5	09VE77 drosophila
5	7	12.1	2424	8	09W206 drosophila
6	6	10.3	33	8	09T219 nicotiana t
7	6	10.3	33	8	09T218 nicotiana t
8	6	10.3	64	4	09GBR1 homo sapien
9	6	10.3	106	10	09ZPW4 arabidopsis
10	6	10.3	112	13	090939 gallus gall
11	6	10.3	126	4	090933 homo sapien
12	6	10.3	127	17	097V04 sulfolobus
13	6	10.3	142	2	09AC17 ehrlichia c
14	6	10.3	142	2	0931A6 bacillus ce
15	6	10.3	144	13	09DET7 bombina ori
16	6	10.3	144	13	P83082 bombina max

17	6	10.3	144	13	P83083	P83083 bombina max
18	6	10.3	144	13	P83081	P83081 bombina max
19	6	10.3	144	13	P83080	P83080 bombina max
20	6	10.3	145	13	P83084	P83084 bombina max
21	6	10.3	158	2	0934J7	0934J7 exiguobacte
22	6	10.3	162	10	09CF9F	09CF9F arabidopsis
23	6	10.3	170	16	0928K0	0928K0 listeria in
24	6	10.3	171	11	063401	063401 rattus norv
25	6	10.3	178	16	09CG00	09CG00 lactococcus
26	6	10.3	186	5	045491	045491 caenorhabd1
27	6	10.3	197	5	061536	061536 drosophila
28	6	10.3	199	5	018251	018251 caenorhabd1
29	6	10.3	201	16	09KB19	09KB19 bacillus ha
30	6	10.3	205	16	045496	045496 bacillus su
31	6	10.3	217	5	09U0V5	09U0V5 leishmania
32	6	10.3	235	5	096193	096193 plasmodium
33	6	10.3	240	4	09NUW9	09NUW9 homo sapien
34	6	10.3	241	16	P94514	P94514 bacillus su
35	6	10.3	247	16	09A8F2	09A8F2 caulobacter
36	6	10.3	248	17	09H113	09H113 thermoplas
37	6	10.3	251	5	09VXC9	09VXC9 drosophila
38	6	10.3	252	2	P71456	P71456 oenococcus
39	6	10.3	252	2	0937E0	0937E0 oenococcus
40	6	10.3	253	16	067075	067075 aquifex ao
41	6	10.3	255	10	09M9Z3	09M9Z3 arabidopsis
42	6	10.3	259	16	09P115	09P115 campylobact
43	6	10.3	262	5	09VRF9	09VRF9 drosophila
44	6	10.3	264	5	020667	020667 caenorhabd1
45	6	10.3	268	16	09W2B9	09W2B9 thermotoga

ALIGNMENTS

RESULT 1
ID 091859 PRELIMINARY: PRT: 433 AA.
AC 091859:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE TRANSCRIPTION FACTOR IIE.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
NCBI_TaxID=8335;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=XENOPUS LAEVIS;
RX MEDLINE=93087200; PubMed=1454543;
RA Ohkuma Y., Hashimoto S., Roeder R.G., Horikoshi M.;
RT "Identification of two large subdomains in TFIIIE-alpha on the basis of
homology between Xenopus and human sequences.";
RL Nucleic Acids Res. 20:5838-5838(1992).
DR EMBL: Z14131; CAI78505.1; -.
DR TRANSFAC: T02165; -.
DR InterPro: IPR002853; TFIIIE_alpha.
DR Pfam: PF02002; TFIIIE_alpha.1.
DR SMART: SM00531; TFIIIE.1.
SQ SEQUENCE 433 AA; 49033 MW; B2C4530BEC2BC00 CRC64;

Query Match 12.1%; Score 7; DB 13; Length 433;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 MFEDLYD 31
|||||
Db 427 MFEDLYD 433

RESULT 2
032519

ID 032519 PRELIMINARY; PRT; 438 AA.
 AC 032519;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE GLYCINE REDUCTASE B.
 GN GRDB.
 OS Eubacterium acidaminophilum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Eubacterium.
 OX NCBI_TaxID=1731;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99191682; PubMed=10091582;
 RA Wagner M., Sontag D., Grimm R., Pich A., Eckerskorn C., Soehling B.,
 RA Andresen J.R.;
 RT "Substrate-specific selenoprotein B of glycine reductase from
 RT Eubacterium acidaminophilum";
 RL Eur. J. Biochem. 260:38-49(1999).
 DR EMBL, Y14275; CAA74651.1;
 DR EMBL, Y14275; CAA74651.1;
 SQ SEQUENCE 438 AA; 47329 MW; 4F411CA0A0DC408 CRC64;

Query Match 12.1%; Score 7; DB 2; Length 438;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 46 KKYAYII 52
 |||||
 DB 129 KKYAYII 135

RESULT 3
 O9VEL0 PRELIMINARY; PRT; 677 AA.
 AC O9VEL0;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CG5319 PROTEIN.
 GN CG5319.

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jajani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Ruidin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 DR EMBL, AE003717; AAF5411.1;
 DR FLYBase: FBgn0038501; CG5319.
 DR InterPro: IPR000210; BTH_P0Z.
 DR Pfam: PF00651; BTH; 1.
 DR SMART: SM00225; BTH; 1.
 DR PROSITE: PS50097; BTH; 1.
 SQ SEQUENCE 677 AA; 72000 MW; CF9FCAC8B09BDFC0 CRC64;

Query Match 12.1%; Score 7; DB 5; Length 677;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 LYCDEIK 16
 |||||
 DB 276 LYCDEIK 282

RESULT 4
 O9VE77 PRELIMINARY; PRT; 844 AA.
 AC O9VE77;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CG7678 PROTEIN.
 GN CG7678.

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jajani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden Klamis I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003722; AAF55550.1; -
 DR FlyBase: FBgn0038613; CG7678.
 DR InterPro: IPR002490; V_Arase_sub_a.
 DR Pfam: PF01496; V_Arase_sub_a; 1.
 SQ SEQUENCE 844 AA; 96652 MW; 672DCA52AD2A284 CRC64;

Query Match 12.1%; Score 7; DB 5; Length 844;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 46 KRYAYII 52
 Db 564 KRYAYII 570

RESULT 5

O9W206 PRELIMINARY; PRT; 2424 AA.

AC O9W206; PRELIMINARY; PRT; 2424 AA.
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE CG13520 PROTEIN.
 GN CG13520.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=BERKELEY;
 RX MEDLINE=201996006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dudin K.J., Evangelista C.C., Ferraz C., Ferrelira S., Fleischmann W.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeigwan C.,
 RA Jajaeli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasio P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Minkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden Klamis I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003458; AAF46894.1; -
 DR FlyBase: FBgn0034771; CG13520.
 SQ SEQUENCE 2424 AA; 283601 MW; FA423F699CB812P0 CRC64;

Query Match 12.1%; Score 7; DB 5; Length 2424;
 Best Local Similarity 100.0%; Pred. No. 83;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 27 EDLYDAK 33
 Db 1206 EDLYDAK 1212

RESULT 6

O9T219 PRELIMINARY; PRT; 33 AA.

AC O9T219;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE RNA BINDING PROTEIN (FRAGMENT).
 OS Nicotiana tabacum (Common tobacco).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OC NCBI_TaxID=4097;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP MEDLINE=93024312; PubMed=1406585;
 RA Mieszcak M., Klahre U., Levy J.H., Goodall G.J., Filipowicz W.;
 RT "Multiple plant RNA binding proteins identified by PCR: expression of
 RT cDNAs encoding RNA binding proteins targeted to chloroplasts in
 RT *Nicotiana plumbaginifolia*.";
 RL Mol. Genet. 234:390-400(1992).
 SQ SEQUENCE 33 AA; 3642 MW; 2F6530107E1CC24B CRC64;

Query Match 10.3%; Score 6; DB 8; Length 33;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 DAKVY 36
 Db 21 DAKVY 26

RESULT 7

O9T218 PRELIMINARY; PRT; 33 AA.

AC O9T218;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE RNA BINDING PROTEIN (FRAGMENT).
 OS Nicotiana tabacum (Common tobacco).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OC NCBI_TaxID=4097;
 RN [1]

```

RP SEQUENCE FROM N.A.
RX MEDLINE=93024312; PubMed=1406585;
RT "Multiple plant RNA binding proteins identified by PCR: expression of
RT cDNAs encoding RNA binding proteins targeted to chloroplasts in
RT Nicotiana glauca."
RL Mol. Gen. Genet. 234:390-400(1992).
SQ SEQUENCE 33 AA; 3656 MW; 01DCA6A4A759DE29 CRC64;

Query Match
Best Local Similarity 10.3%; Score 6; DB 8; Length 33;
Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 DAKVY 36
DB 21 DAKVY 26

RESULT 8
Q96RF1 PRELIMINARY; PRT; 64 AA.
ID Q96RF1
AC Q96RF1
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HbJ1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Rhodes C.H., Gall K.M., Little R., Braunschweiler K., Park J.P.;
RT "NOE3: a novel olfactomedin/noelin/pancortin homolog identified near
RT an endymoma-associated translocation breakpoint."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF395440; AAK73122.1;
SQ SEQUENCE 64 AA; 7298 MW; 3CAE1FF08C6E8B6 CRC64;

Query Match
Best Local Similarity 10.3%; Score 6; DB 4; Length 64;
Pred. No. 44;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 EDLYDA 32
DB 19 EDLYDA 24

RESULT 9
Q92PW4 PRELIMINARY; PRT; 106 AA.
ID Q92PW4
AC Q92PW4
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE AT2G18320 PROTEIN.
GN AT2G18320.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurossids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA.
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanden S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Cressy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;

```

```

RP "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana."
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA.
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006439; AAD15505.1;
SQ SEQUENCE 106 AA; 12023 MW; A6873165916B150C CRC64;

Query Match
Best Local Similarity 10.3%; Score 6; DB 10; Length 106;
Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 KYAVII 52
DB 71 KYAVII 76

RESULT 10
Q90939 PRELIMINARY; PRT; 112 AA.
ID Q90939
AC Q90939
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)
DE P120 (FRAGMENT).
OS Gallus gallus (chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Sawada K., Agata K., Eguchi G.;
RT "Characterization of cDNA clones derived from chicken lens fibers."
RL Int. J. Dev. Biol. 40:531-535(1996).
DR EMBL; D49741; BAA08575.1;
SQ SEQUENCE 112 AA; 12450 MW; D40BDFAI40EDB08E CRC64;

Query Match
Best Local Similarity 10.3%; Score 6; DB 13; Length 112;
Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ERKXT 7
DB 92 ERKXT 97

RESULT 11
Q9UGT3 PRELIMINARY; PRT; 126 AA.
ID Q9UGT3
AC Q9UGT3
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE D671014.1 (CGI-56 SIMILAR TO C.ELEGANS PROTEIN T21012.4.)
GN D671014.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Matthews L.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031595; CAB63068.1;
SQ InterPro; IPR001715; Calponin_hom.

```

DR Pfam; PF00307; CH; 1.
DR SMART; SM00033; CH; 1.
DR PROSITE; PS0021; CH; 1.
FT NON_TER
SQ SEQUENCE 126 AA; 14603 MW; EC2DC72F31EC8A4 CRC64;

Query Match
Best Local Similarity 10.3%; Score 6; DB 4; Length 126;
100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 41 YNLFK 46
|||||
DB 116 YNLFK 121

RESULT 12

O97V04 PRELIMINARY; PRT; 127 AA.

AC O97V04;
DT 01-OCT-2001 (TREMblrel. 18, Created)
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
DE 01-OCT-2001 (TREMblrel. 18, Last annotation update)
DE HYPOTHEICAL PROTEIN SS02833.
GN SS02833.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=2133296; PubMed=11427726;
RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aways M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moers A., Eruso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; AE006877; AA02943.1; -;
DR InterPro; IPR000130; 2n_MTPeptide.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 127 AA; 14923 MW; 19B865F700A04B3 CRC64;

Query Match
Best Local Similarity 10.3%; Score 6; DB 17; Length 127;
100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 KVVYSY 38
|||||
DB 32 KVVYSY 37

RESULT 13

O9AC17 PRELIMINARY; PRT; 142 AA.

AC O9AC17;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
DE HYPOTHEICAL 15.2 KDA PROTEIN.
OS Ehrlichia chaffeensis.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Ehrlichiaeae; Ehrlichia.
OX NCBI_TaxID=945;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ARKANSAS;
RX MEDLINE=21153566; PubMed=11254561;
RA Ohashi N., Rikihisa Y., Unver A.;
RT "Analysis of Transcriptionally Active Gene Clusters of Major Outer

RT Membrane Protein Multigene Family in Ehrlichia canis and E.
RT chaffeensis.";
RL Infect. Immun. 69:2083-2091(2001).
DR EMBL; U72291; AK28677.1; -;
DR InterPro; IPR000508; Peptidase_S26.
DR PROSITE; PS00761; SPASE_1_3; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 142 AA; 15246 MW; CDF9C9BF1565F63 CRC64;

Query Match
Best Local Similarity 10.3%; Score 6; DB 2; Length 142;
100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 50 YIIIEYI 55
|||||
DB 60 YIIIEYI 65

RESULT 14

O93LA6 PRELIMINARY; PRT; 142 AA.

AC O93LA6;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE TRANSPOSASE (FRAGMENT).
GN TNP.A.
OS Bacillus cereus.
OC Plasmid pRI46.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1396;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RC607; TRANSPOSON=TN5084;
RX MEDLINE=21339354; PubMed=11446519;
RA Bodanova E., Minakhin L., Baas I., Hobman J., Volodin A.,
RA Nikiforov V.;
RT "Class II broad-spectrum mercury resistance transposons in Gram-
RT positive bacteria from natural environments.";
RL Res. Microbiol. 152:503-514(2001).
DR EMBL; Y17748; CAC41962.1; -;
KW Plasmid.
FT NON_TER
SQ SEQUENCE 142 AA; 16788 MW; 557733EECC7B3BDC CRC64;

Query Match
Best Local Similarity 10.3%; Score 6; DB 2; Length 142;
100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 50 YIIIEYI 55
|||||
DB 76 YIIIEYI 81

RESULT 15

O9DET7 PRELIMINARY; PRT; 144 AA.

AC O9DET7;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE BOMBININ-LIKE PEPTIDE 7, BPL-7 PRECURSOR.
GN BLP-7.
OS Bombina orientalis (Oriental fire-bellied toad).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Archeobatrachia; Bombinatoridae; Bombina.
OX NCBI_TaxID=8346;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=20545783; PubMed=11090922;
RA Miele R., Botto M., Fiocco D., Barra D., Simeone M.;

RT "Sequence of a gene from Bombina orientalis coding for the
 RT antimicrobial peptide BLP-7."
 RL Peptides 21:1681-1686(2000).
 DR EMBL: AJ298827; CAC11122.1; -.
 KW Signal.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 44 70 BOMBININ-LIKE PEPTIDE 7, BLP-7.
 FT CHAIN 124 144 GENE-DERIVED BOMBININ H-TYPE PEPTIDE 2,
 FT GH-2.
 SQ SEQUENCE 144 AA; 16053 MW; 4ECEFCA81D62764B CRC64;

Query Match 10.3%; Score 6; DB 13; Length 144;
 Best Local Similarity 100.0%; Pred. No. 88;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 NLETKK 47
 |||||
 DB 115 NLETKK 120

RESULT 16
 P83082 PRELIMINARY; PRT; 144 AA.

ID P83082;
 AC P83082; 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE MAXIMIN 3 PRECURSOR [CONTAINS: MAXIMIN-3].
 OS Bombina maxima.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Archeobatrachia; Bombinatoridae; Bombina.
 OX NCBI_TaxID=161274;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Y., Lai R., Zheng Y., Lee W.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SKIN SECRETION;
 RA Lai R., Zheng Y.-T., Shen J.-H., Liu H., Lee W.-H., Tang S.-Z.,
 RA Zhang Y.;
 RT "Antimicrobial peptides from the skin secretion of Chinese red belly
 RT toad Bombina maxima."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: HAS ANTIMICROBIAL ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE BOMBININ FAMILY.
 DR EMBL: AF378906; AAK63256.1; -.
 KW Signal; Antibiotic.
 KM SIGNAL 1 18 POTENTIAL.
 FT PEPTIDE 44 70 MAXIMIN-3.
 FT PEPTIDE 144 AA; 16099 MW; D1AA7B247ID9BBBC CRC64;
 SQ SEQUENCE

Query Match 10.3%; Score 6; DB 13; Length 144;
 Best Local Similarity 100.0%; Pred. No. 88;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 NLETKK 47
 |||||
 DB 115 NLETKK 120

RESULT 17
 P83083 PRELIMINARY; PRT; 144 AA.

ID P83083;
 AC P83083; 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE MAXIMIN 4 PRECURSOR [CONTAINS: MAXIMIN-4].
 OS Bombina maxima.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Archeobatrachia; Bombinatoridae; Bombina.

OX NCBI_TaxID=161274;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SKIN SECRETION;
 RA Lai R., Zheng Y.-T., Shen J.-H., Liu H., Lee W.-H., Tang S.-Z.,
 RA Zhang Y.;
 RT "Antimicrobial peptides from the skin secretion of Chinese red belly
 RT toad Bombina maxima."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 44-70, AND CHARACTERIZATION.
 RC TISSUE-SKIN SECRETION;
 RA Chen T.B., McClean S., Orr D.F., Bjorson A.J., Rao P.F., Shaw C.;
 RT "Isolation and structural characterisation of antimicrobial peptides
 RT from the venom of the Chinese large-webbed bell toad (Bombina
 RT maxima)."
 RL Submitted (JUL-2001) to the SWISS-PROT data bank.
 CC -1- FUNCTION: HAS ANTIMICROBIAL ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- TISSUE SPECIFICITY: SKIN.
 CC -1- SIMILARITY: BELONGS TO THE BOMBININ FAMILY.
 DR EMBL: AF378905; AAK63255.1; -.
 KW Signal; Antibiotic.
 KM SIGNAL 1 18 POTENTIAL.
 FT PEPTIDE 44 70 MAXIMIN-2.
 FT PEPTIDE 144 AA; 15960 MW; 6124AV970C63CE7 CRC64;
 SQ SEQUENCE

Query Match 10.3%; Score 6; DB 13; Length 144;
 Best Local Similarity 100.0%; Pred. No. 88;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 NLETKK 47
 |||||
 DB 115 NLETKK 120

RESULT 18
 P83081 PRELIMINARY; PRT; 144 AA.

ID P83081;
 AC P83081; 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE MAXIMIN 2 PRECURSOR [CONTAINS: MAXIMIN-2].
 OS Bombina maxima.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Archeobatrachia; Bombinatoridae; Bombina.
 OX NCBI_TaxID=161274;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lai R., Zheng Y.-T., Shen J.-H., Liu H., Lee W.-H., Tang S.-Z.,
 RA Zhang Y.;
 RT "Antimicrobial peptides from the skin secretion of Chinese red belly
 RT toad Bombina maxima."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 44-70, AND CHARACTERIZATION.
 RC TISSUE-SKIN SECRETION;
 RA Chen T.B., McClean S., Orr D.F., Bjorson A.J., Rao P.F., Shaw C.;
 RT "Isolation and structural characterisation of antimicrobial peptides
 RT from the venom of the Chinese large-webbed bell toad (Bombina
 RT maxima)."
 RL Submitted (JUL-2001) to the SWISS-PROT data bank.
 CC -1- FUNCTION: HAS ANTIMICROBIAL ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- TISSUE SPECIFICITY: SKIN.
 CC -1- SIMILARITY: BELONGS TO THE BOMBININ FAMILY.
 DR EMBL: AF378905; AAK63255.1; -.
 KW Signal; Antibiotic.
 KM SIGNAL 1 18 POTENTIAL.
 FT PEPTIDE 44 70 MAXIMIN-2.
 FT PEPTIDE 144 AA; 15960 MW; 6124AV970C63CE7 CRC64;
 SQ SEQUENCE

Query Match 10.3%; Score 6; DB 13; Length 144;
 Best Local Similarity 100.0%; Pred. No. 88;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 NLETKK 47
 |||||
 Db 115 NLETKK 120

RESULT 19

ID P83080 PRELIMINARY; PRT; 144 AA.
 AC P83080;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DE MAXIMIN 1 PRECURSOR [CONTAINS: MAXIMIN-1].
 OS Bombina maxima.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Archeobatrachia; Bombinatoridae; Bombina.
 OX NCBI_TaxID=161274;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SKIN SECRETION;
 RA Lai R., Zheng Y.-T., Shen J.-H., Liu H., Lee W.-H., Tang S.-Z.,
 Zhang Y.;
 RT "Antimicrobial peptides from the skin secretion of Chinese red belly
 toad Bombina maxima.";
 RT Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 44-70, AND CHARACTERIZATION.
 RC TISSUE-SKIN SECRETION;
 RA Chen T.B., McClean S., Orr D.F., Bjourson A.J., Rao P.F., Shaw C.;
 RT "Isolation and structural characterisation of antimicrobial peptides
 from the venom of the Chinese large-webbed bell toad (Bombina
 maxima).";
 RT Submitted (JUL-2001) to the SWISS-PROT data bank.
 CC -1- FUNCTION: HAS ANTIMICROBIAL ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- TISSUE SPECIFICITY: SKIN.
 CC -1- SIMILARITY: BELONGS TO THE BOMBININ FAMILY.
 DR EMBL; AF378904; AAK63254.1; -.
 KW Signal; Antibiotic.
 FT SIGNAL. 1 18 POTENTIAL.
 FT PEPTIDE 44 70 MAXIMIN-5.
 SQ SEQUENCE 144 AA; 15976 MW; F770F6870543020E CRC64;

Query Match 10.3%; Score 6; DB 13; Length 144;
 Best Local Similarity 100.0%; Pred. No. 88;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 NLETKK 47
 |||||
 Db 115 NLETKK 120

RESULT 20

ID P83084 PRELIMINARY; PRT; 145 AA.
 AC P83084;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DE MAXIMIN 5 PRECURSOR [CONTAINS: MAXIMIN-5].
 OS Bombina maxima.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Archeobatrachia; Bombinatoridae; Bombina.
 OX NCBI_TaxID=161274;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Y., Lai R., Lee W.;
 RT Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.

RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SKIN SECRETION;
 RA Lai R., Zheng Y.-T., Shen J.-H., Liu H., Lee W.-H., Tang S.-Z.,
 Zhang Y.;
 RT "Antimicrobial peptides from the skin secretion of Chinese red belly
 toad Bombina maxima.";
 RT Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 44-70, AND CHARACTERIZATION.
 RC TISSUE-SKIN SECRETION;
 RA Chen T.B., McClean S., Orr D.F., Bjourson A.J., Rao P.F., Shaw C.;
 RT "Isolation and structural characterisation of antimicrobial peptides
 from the venom of the Chinese large-webbed bell toad (Bombina
 maxima).";
 RT Submitted (JUL-2001) to the SWISS-PROT data bank.
 CC -1- FUNCTION: HAS ANTIMICROBIAL ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- TISSUE SPECIFICITY: SKIN.
 CC -1- SIMILARITY: BELONGS TO THE BOMBININ FAMILY.
 DR EMBL; AF378908; AAK63258.1; -.
 KW Signal; Antibiotic.
 FT SIGNAL. 1 18 POTENTIAL.
 FT PEPTIDE 44 70 MAXIMIN-5.
 SQ SEQUENCE 145 AA; 16338 MW; CD7A93483196F7E3 CRC64;

Query Match 10.3%; Score 6; DB 13; Length 145;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 NLETKK 47
 |||||
 Db 116 NLETKK 121

RESULT 21

ID O934J7 PRELIMINARY; PRT; 158 AA.
 AC O934J7;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE TRANSPOSASE (FRAGMENT).
 GN TNPA.
 OS Exiguobacterium sp.
 OG Plasmid pKLH3.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Exiguobacterium.
 OX NCBI_TaxID=44751;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TG38-2B; TRANSPOSON-TNS085;
 RX MEDLINE=21339354; PubMed=11446519;
 RA Bogdanova E., Minakhin L., Bass I., Hobman J., Volodina A.,
 Nikiforov V.;
 RT "Class II broad-spectrum mercury resistance transposons in Gram-
 positive bacteria from natural environments.";
 RT Res. Microbiol. 152:503-514 (2001).
 DR EMBL; Y17750; CAC41997.1; -.
 KW Plasmid.
 FT NON-TER. 158
 SQ SEQUENCE 158 AA; 18708 MW; B04E51BF490FF2DA CRC64;

Query Match 10.3%; Score 6; DB 2; Length 158;
 Best Local Similarity 100.0%; Pred. No. 96;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 50 YIIEYI 55
 |||||
 Db 76 YIIEYI 81

RESULT 22

09CGP9 ID 09CGP9 PRELIMINARY; PRT; 162 AA.
 AC 09CGP9;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE HYPOTHEICAL 18.9 KDA PROTEIN.
 GN T22E19.13
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 Gall J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huzar L.,
 Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani L.,
 Miltischer J., Miranda M., Nguyen M., Nierman M.C., Osborne B.I.,
 Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 Salgado H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 Utechtack T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana."
 RL Nature 408:816-820(2000).
 DR EMBL: AC016447; AAC52603.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 162 AA; 18940 MW; A38961D280B68FE CRC64;
 Query Match 10.3%; Score 6; DB 10; Length 162;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 51 IIEYIK 56
 |||||
 Db 87 IIEYIK 92
 RESULT 23
 Q928K0 ID 0928K0 PRELIMINARY; PRT; 170 AA.
 AC 0928K0;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE LIN2532. PROTEIN.
 GN LIN2532.
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Listeria.
 OX NCBI_TaxID=1642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIP 11262 / SEROVAR 6A;
 RX PubMed=11679669;
 RA Glaeser P., Frangoul L., Buchrieser C., Rusniok C., Amend A.,
 Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 Charbit A., Chetouani F., Couve E., de Darvar A., Dehoux P.,
 Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 Entian K.-D., Esibi H., Garcia-del Portillo F., Garrido P.,
 Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,

RA Madueno E., Maitounam A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
 RT "Comparative genomics of Listeria species."
 RL Science 294:849-852(2001).
 DR EMBL: AL596172; CAC97759.1; -
 DR ListList; LIN02532; -
 KW Complete proteome.
 SQ SEQUENCE 170 AA; 20178 MW; C6DFA34727FD8639 CRC64;
 Query Match 10.3%; Score 6; DB 16; Length 170;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 43 LFTKRY 48
 |||||
 Db 127 LFTKRY 132
 RESULT 24
 Q63401 ID 063401 PRELIMINARY; PRT; 171 AA.
 AC 063401;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE (CLONE REM2) ORF (FRAGMENT).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HOLZMAN; TISSUE=BRAIN;
 RX MEDLINE=96235155; PubMed=8642059;
 RA Asakura K., Pogulis R.J., Pease L.R., Rodriguez M.;
 RT "A monoclonal autoantibody which promotes central nervous system
 remyelination is highly polyreactive to multiple known and novel
 RT antigens."
 RL J. Neuroimmunol. 65:11-19(1996).
 DR EMBL: LA1684; AAB05842.1; -
 DT NON_TER 1 171
 FT NON_TER 1 171
 SQ SEQUENCE 171 AA; 18947 MW; 36576590383E0581 CRC64;
 Query Match 10.3%; Score 6; DB 11; Length 171;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 30 YDARKV 35
 |||||
 Db 38 YDARKV 43
 RESULT 25
 Q9CG00 ID 09CG00 PRELIMINARY; PRT; 178 AA.
 AC 09CG00;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE UNKNOWN PROTEIN.
 GN YNFE.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=111403;
 RX MEDLINE=21235186; PubMed=11337471;
 RA Bolotin A., Winkler P., Manger S., Jaillon O., Malarne K.,

RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium *Lactococcus*
RI *lactis* ssp. *lactis* IL1403.";
RL Genome Res. 11:731-753(2001).
DR EMBL; AE006363; AAK05409.1; -.
KM Complete Proteome.
SQ SEQUENCE 178 AA; 21344 MW; E7F0004422467B51 CRC64;

Query Match 10.3%; Score 6; DB 16; Length 178;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 LEFTKY 48
|||||
DB 136 LEFTKY 141

RESULT 26
O45491 PRELIMINARY; PRT; 186 AA.

ID 045491
AC 045491;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE F38C2.5 PROTEIN.
GN F38C2.5.
OS *Caenorhabditis elegans*.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; *Caenorhabditis*.
OX NCBI_TaxID=6239;

RN [1]
RP SEQUENCE FROM N.A.

RA Mortimore B.J.;
RN Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.

RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;

RA none;
RT "Genome sequence of the nematode *C. elegans*: A platform for
RT investigating biology.";

RL Science 282:2012-2018(1998).
DR EMBL; Z82267; CAB05191.1; -.
DR InterPro: IPR000571; Zf-CCCH.

DR Pfam; PF00642; Zf-CCCH; 2.
DR SMART: SM00356; Znf_C3H1; 2.

SO SEQUENCE 186 AA; 21118 MW; 4A9C16B15641853F CRC64;

Query Match 10.3%; Score 6; DB 5; Length 186;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KYKTVL 9
|||||
DB 115 KYKTVL 120

RESULT 27

O61536 PRELIMINARY; PRT; 197 AA.

ID 061536
AC 061536;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TRYPSTIN (FRAGMENT).
GN TRY.

OS *Drosophila heteroneura* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; *Drosophila*.
OX NCBI_TaxID=32382;

RN [1]
RP SEQUENCE FROM N.A.
RA Davis T., Kurihara J., Yamamoto D.;

RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF049235; AAC05141.1; -.
DR HSSP; P00761; 1EPT.
DR MEROPS; S01.0PA; -.
DR FlyBase; FBgn0025437; Dhet\TRY.

DR InterPro: IPR001254; Trypsin.
DR Pfam; PF00089; trypsin; 1.
DR SMART; SM00020; Tryp_Spc; 1.

DR PROSITE; PS00240; TRYPSTIN_DOM; 1.
DR PROSITE; PS00134; TRYPSTIN_HIS; UNKNOWN_1.

DR Hydrolase; Serine protease.
KM NON_TER
FT NON_TER
SQ SEQUENCE 197 AA; 21199 MW; 6CE8CE5BD97F7F94 CRC64;

Query Match 10.3%; Score 6; DB 5; Length 197;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 PHQISM 25
|||||
DB 42 PHQISM 47

RESULT 28
O18251 PRELIMINARY; PRT; 199 AA.

ID 018251
AC 018251;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Y57G11C.25 PROTEIN.
GN Y57G11C.25.
OS *Caenorhabditis elegans*.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; *Caenorhabditis*.
OX NCBI_TaxID=6239;

RN [1]
RP SEQUENCE FROM N.A.

RA McMurray A.A.;
RN Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.

RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;

RA none;
RT "Genome sequence of the nematode *C. elegans*: A platform for
RT investigating biology.";

RL Science 282:2012-2018(1998).
DR EMBL; Z99281; CAB16528.1; -.
DR InterPro: IPR000571; Zf-CCCH.

DR Pfam; PF00642; Zf-CCCH; 2.
DR SMART; SM00356; Znf_C3H1; 2.

SO SEQUENCE 199 AA; 22786 MW; B1839B24922004EE CRC64;

Query Match 10.3%; Score 6; DB 5; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KYKTVL 9
|||||
DB 111 KYKTVL 116

RESULT 29

O9KBJ9 PRELIMINARY; PRT; 201 AA.

ID 09KBJ9
AC 09KBJ9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE BH1927 PROTEIN.

GN BH1927.
OS *Bacillus halodurans*.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL; AP001513; BAB05646.1; -;
 KW Complete proteome.
 SQ SEQUENCE 201 AA; 22898 MW; DAB8A10A5E76A951 CRC64;
 Query Match 10.3%; Score 6; DB 16; Length 201;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 44 FTKRYA 49
 Db 195 FTKRYA 200
 RESULT 30
 Q45496 PRELIMINARY; PRT; 205 AA.
 AC Q45496;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DE 01-NOV-1998 (TRENBLREL. 08, Last annotation update)
 DE HYPOTHEORETICAL 23.9 KDA PROTEIN.
 GN YKRC OR YKVA.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-1168;
 RX MEDLINE=90368558; PubMed=1697575;
 RA Hemila H., Palva A., Paulin L., Arvidson S., Palva I.;
 RT "Secretory S complex of Bacillus subtilis: sequence analysis and
 RT identity to pyruvate dehydrogenase.";
 RL J. Bacteriol. 172:5052-5063(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-1168;
 RX MEDLINE=97124187; PubMed=8969500;
 RA Winters P., Caldwell R., Enfield L., Ferrari E.;
 RT "The ampS-nprE (124 degrees-127 degrees) region of the Bacillus
 RT subtilis 168 chromosome: sequencing of a 27 kb segment and
 RT identification of several genes in the area.";
 RL Microbiology 142:0-0(0).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-1168;
 RX MEDLINE=97144523; PubMed=8990290;
 RA Henriques A.O., Bryan E.M., Beall B.W., Moran C.P. Jr.;
 RT "cseJ, cseG, and csk22 are new members of mother-cell-specific
 RT sporulation regulons in Bacillus subtilis.";
 RL J. Bacteriol. 179:389-398(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Boriss L., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell R., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,

SA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 SA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 SA Fritz C., Fujita M., Fujita Y., Fuma S., Gallazzi A., Galleron N.,
 SA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 SA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 SA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 SA Joris B., Karamata D., Kasahara Y., Kieffer-Blanchard M., Klein C.,
 SA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 SA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 SA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 SA Medina N., Meliado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 SA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 SA Parro V., Pohl T.M., Portetelle D., Portwollik S., Prescott A.M.,
 SA Priesean E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 SA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
 SA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 SA Sekiguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,
 SA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 SA Takeuchi M., Tanakoshi A., Tanaka T., Terstra P., Tognoni A.,
 SA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 SA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
 SA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 SA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zumbstein A.;
 RT "The complete genome sequence of the gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256(1997).
 DR EMBL; AF012285; AAC24931.1; -;
 DR EMBL; Z99111; CAB1330.1; -;
 KW Hypothetical protein; Complete proteome.
 RT CONFICT 77
 GO SEQUENCE 205 AA; 23912 MW; 5C9A604C240E91C8 CRC64;
 Query Match 10.3%; Score 6; DB 16; Length 205;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 27 EDLYDA 32
 Db 192 EDLYDA 197
 RESULT 31
 Q900V5 PRELIMINARY; PRT; 217 AA.
 AC Q900V5;
 DT 01-MAY-2000 (TRENBLREL. 13, Created)
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
 DE POSSIBLE DIMPI HOMOLOG.
 GN L7836.07.
 OS Leishmania major.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 NCBI_TaxID=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FRIEDLIN;
 RA Oliver K., Harris D., Ivens A.C., Lawson D., Quail M.,
 RA Rajandream M.A., Barrell B.G.;
 RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FRIEDLIN;
 RX MEDLINE=98146435; PubMed=9477341;
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
 RA Smith D.F.;
 RT "A physical map of the Leishmania major Friedlin genome.";
 RL Genome Res. 8:135-145(1998).
 DR EMBL; AL117263; CAB5382.1; -;
 DR InterPro; IPR004123; DIM1.
 DR Pfam; PF02966; DIM1; 1.
 GO SEQUENCE 217 AA; 24255 MW; F90138C9866961D4 CRC64;
 Query Match 10.3%; Score 6; DB 5; Length 217;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 FEDLYD 31
Db 180 FEDLYD 185

RESULT 32

096193 ID 096193 PRELIMINARY; PRT; 235 AA.
AC 096193;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RAB GTPASE.
GN PFE0500C.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99021743; PubMed-9804551.
RA Gardner M.J., Tettelin H., Carucci D.J., Cummings L.M., Aravind L., Koonin E.V., Shalim S., Mason T., Yu K., Fujii C., Pederson J., Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pertea M., Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O., Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum."
RL Science 282:1126-1132(1998).
CC -1 SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL: AE001399; AAC71889.1; -
DR HSSP: P36017; IER0.
DR InterPro: IPR003579; Rab.
DR InterPro: IPR001806; Ras_trnsfmrng.
DR InterPro: IPR003575; Small_GTPase.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRNG.
DR SMART: SM00175; RAB; 1.
DR SMART: SM00010; small_GTPase; 1.
DR GTP-binding; Lipoprotein.
KW SEQUENCE 235 AA; 26671 MW; 448A6813CDB99A84 CRC64;
SQ

Query Match 10.3%; Score 6; DB 5; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKTVL 10
Db 7 YKTVL 12

RESULT 33

09NMW9 ID 09NMW9 PRELIMINARY; PRT; 240 AA.
AC 09NMW9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE CDNA FLJ11088 FTS, CLONE PLACE1005287, WEAKLY SIMILAR TO INNER CENTROMERE PROTEIN.
DE Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PLACENTA;
RA Itoigai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y., Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RA

RT "NEDO human cDNA sequencing project."
RL Submitted (FE8-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK001950; BAA9195.1; -
SQ SEQUENCE 240 AA; 27958 MW; AD9D4C347990CB66 CRC64;

Query Match 10.3%; Score 6; DB 4; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 IEYKE 57
Db 196 IEYKE 201

RESULT 34

P94514 ID P94514 PRELIMINARY; PRT; 241 AA.
AC P94514;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE AUTOLYSIN RESPONSE REGULATOR.
GN LYPT.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
CC Bacillus/Staphylococcus group; Bacillus.
NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97217425; PubMed-9063446;
RA Goethel S.F., Schmid R., Wipat A., Carter N.M., Emerson P.T., Harwood C.R., Marbuhl M.A.;
RT "An internal FK506-binding domain is the catalytic core of the prolyl isomerase activity associated with the Bacillus subtilis trigger factor."
RL Eur. J. Biochem. 244:59-65(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-168;
RX MEDLINE-98044033; PubMed-9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertoletti M.G., Bessieres P., Bolotin A., Borchert S., Borries R., Boursier L., Brans A., Braun M., Brigelli S.C., Bron S., Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoft A., Ehtilich S.D., Emerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G., Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M., Klein C., Kobayashi Y., Koeltter P., Koningsstein G., Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Patro V., Pohl T.M., Portetelle D., Rappoport G., Prescott A.M., Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadleir Y., Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassartoli A., Viari A., Wambuit R., Wedler E., Wedler H., Welleneger T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the gram-positive bacterium Bacillus subtilis."
RL Nature 390:249-256(1997).
RN [3]
RP SEQUENCE FROM N.A.

RC STRAIN-168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATOR COMPONENTS OF SENSOR TRANSDUCTION SYSTEMS.
DR EMBL: Z75208; CA99611.1; -;
DR EMBL: Z99118; CAB14852.1; -;
DR HSSP: Q56312; 1TMV.
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF00072; response_reg.1.
DR SMART: SM00448; REC.1.
KW Complete proteome; Phosphorylation; Sensory transduction.
SQ SEQUENCE 241 AA; 27834 MW; 315917932996089F CRC64;

Query Match 10.3%; Score 6; DB 16; Length 241;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 EYKEI 58
DB 202 EYKEI 207

RESULT 35
ID Q9ABF2 PRELIMINARY; PRT; 247 AA.
AC Q9ABF2;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CYTOCHROME C OXIDASE, CCOO SUBUNIT.
GN CCI402.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=69394;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE-21173698; PubMed-11259647;
RA Niernan W.C., Feldblyum T.V., Laud M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohia N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Sait J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL: AE005815; AKR23383.1; -;
DR TIGR: CCI402; -;
DR InterPro: IPR000345; CytC_heme_bind.
DR InterPro: IPR003468; F1XO.1
DR Pfam: PF02433; F1XO.1
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 247 AA; 27975 MW; 79E8B0FAF8A6A06A CRC64;

Query Match 10.3%; Score 6; DB 16; Length 247;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 YDAKV 35
DB 200 YDAKV 205

RESULT 36
ID Q9HIL3 PRELIMINARY; PRT; 248 AA.
AC Q9HIL3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PROBABLE DIPEPTIDE TRANSPORT SYSTEM, ATP-BINDING PROTEIN DPPF.
IN TAI326.
OS Thermoplasma acidophilum.
CC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmaceae;
CC Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DSM 1728;
RX MEDLINE-20479972; PubMed-11029001;
RA Ruepp A., Gräml W., Santos Martinez M.-L., Koretke K.K., Volker C.,
RA Wewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermophilic scavenger Thermoplasma
acidophilum.";
RL Nature 407:508-513(2000).
DR EMBL: AL445067; CAC12447.1; -;
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00005; ABC_tran.1.
DR SMART: SM00382; AAA.1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 248 AA; 28424 MW; 691B064069F41DB3 CRC64;

Query Match 10.3%; Score 6; DB 17; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 FPHQIS 24
DB 134 FPHQIS 139

RESULT 37
ID Q9VXC9 PRELIMINARY; PRT; 251 AA.
AC Q9VXC9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CG9676 PROTEIN.
GN CG9676.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abjil J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
RA Beson R.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burkis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrelia S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mervlov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Piltman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
 RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 CC EMBL: AE003503; AAF48647.1; -.
 DR HSSP: P00761; 1EPT.
 DR MEROPS: S01.URA.1.
 DR Flybase: FBgn0030774; CG9675.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Trypsin.
 DR Pfam: PF00089; trypsin.1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; TRYP-SPEC.1.
 DR PROSITE: PS00240; TRYPSIN_DOM.1.
 DR PROSITE: PS00134; TRYPSIN_HIS.1.
 DR Hydrolase: Serine protease.
 SQ SEQUENCE 251 AA; 26666 MW; 2248038D6DCDF809 CRC64;

Query Match 10.3%; Score 6; DB 5; Length 251;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 PPROIS 24
 DB 39 PPROIS 44

RESULT 38
 P71456 PRELIMINARY; PRT; 252 AA.
 AC P71456;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ORF2 PROTEIN.
 OS *Oenococcus oeni* (Leuconostoc oeni).
 OC Plasmid p0332.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
 OC *Oenococcus*.
 OX NCBI_TaxID=1247;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-B0632;
 RX MEDLINE=97092387; PubMed=8938052;
 RA Brito L., Vieira G., Santos M.A., Paveia H.;
 RT "Nucleotide Sequence Analysis of p0332, a Cryptic Plasmid from
 RT *Leuconostoc oeni*.";
 RL Plasmid 36:49-54(1996).
 DR EMBL: X86402; CAA60155.1; -.
 KW Plasmid.
 SO SEQUENCE 252 AA; 29553 MW; 398EF169691A9F96 CRC64;

Query Match 10.3%; Score 6; DB 2; Length 252;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 3 RRYKTV 8
 DB 11111

DB 69 RRYKTV 74
 RESULT 39
 ID Q937E0 PRELIMINARY; PRT; 252 AA.
 AC Q937E0;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE RECOMBINATION PROTEIN.
 GN PRE.
 OS *Oenococcus oeni* (Leuconostoc oeni).
 OC Plasmid pRS2.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
 OC *Oenococcus*.
 OX NCBI_TaxID=1247;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mesas J.M., Rodriguez M.C., Alegre M.T.;
 RT "Nucleotide sequence analysis of pRS2 and pRS3, two small cryptic
 RT plasmids from an *Oenococcus oeni* strain.";
 RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AJ310613; CAC67506.1; -.
 KW Plasmid.
 SQ SEQUENCE 252 AA; 29337 MW; 0AE79405E509C118 CRC64;

Query Match 10.3%; Score 6; DB 2; Length 252;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 RRYKTV 8
 DB 69 RRYKTV 74

RESULT 40
 ID Q67075 PRELIMINARY; PRT; 253 AA.
 AC Q67075;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE HYPOTHETICAL 29.0 KDA PROTEIN.
 GN AQ_933.
 OS *Aquifex aeolicus*.
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VF5;
 RX MEDLINE=96196666; PubMed=9537320;
 RA Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjey M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swenson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium *Aquifex*
 RT *aeolicus*.";
 RL Nature 392:353-358(1998).
 DR EMBL: AE000714; AAC07040.1; -.
 DR InterPro: IPR001932; PP2C_domain.
 DR InterPro: IPR003643; Putat_PPase.
 DR Pfam: PF00481; PP2C.1.
 DR ProDom: PD006823; Putat_PPase.1.
 DR SMART: SM00332; PP2C; 1.
 DR SMART: SM00331; PP2C-SIG.1.
 DR Hypothetical protein: Complete proteome.
 KW Hypothetical protein: Complete proteome.
 SO SEQUENCE 253 AA; 29015 MW; B8E0D6CD/F92D890 CRC64;

Query Match 10.3%; Score 6; DB 16; Length 253;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 53 EYIKI 58

Db 82 EYKEI 87

RESULT 41

09M923 PRELIMINARY; PRT; 255 AA.
AC 09M923;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
DE F4H5.5 PROTEIN.
GN F4H5.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids; II: Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Federici N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
RA Gonzalez A., Kremenetskaia I., Kim C., Ienz C., Li J., Liu S.,
RA Luoro S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,
RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC011001; AAF63135.1; -
DR InterPro: IPR000620; DUF6.
DR InterPro: IPR000623; SHK_kinase.
DR Pfam: PF00892; DUF6; 1.
DR PROSITE: PS01128; SHIKIMATE_KINASE; UNKNOWN.1.
SQ SEQUENCE 255 AA; 2797 MM; 2C3F157FDA34FE79 CRC64;

Query Match 10.3%; Score 6; DB 10; Length 255;
Best Local Similarity 100.0%; Pred. No. 14e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 VVSYX 39
Db 189 VVSYX 194

RESULT 42

09P015 PRELIMINARY; PRT; 259 AA.
AC 09P015;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE PERIPLASMIC PROTEIN.
GN C7011.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-NCTC 11168;
RX MEDLINE-20150912; PubMed-10688204;
RA Parkhill J., Wren B.W., Mungall K., Kelsey J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagers K., Karlyshew A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrett B.G.;
RL "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RT Nature 403:665-668(2000).
RU EMBL: AL139074; CAB72595.1; -
SQ SEQUENCE 259 AA; 29294 MM; AC401B4070B1D36F CRC64;

Query Match 10.3%; Score 6; DB 16; Length 259;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 29 LYDAKY 34
Db 218 LYDAKY 223

RESULT 43

09VRF9 PRELIMINARY; PRT; 262 AA.
AC 09VRF9;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE CG1497 PROTEIN.
GN CG1497.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaral P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Burton R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ipeywan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasto P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshirei A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock K.C., Wu D., Yang S., Yao Q.A.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of Drosophila melanogaster.";
RT Science 287:2185-2195(2000).
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.

EMBL: AE003569; AAF50842.1; -
HSSP: P20160; 1A7S.
DR MEROPS: S01.0PA; -
DR FLYbase: FBgn0031167; CG1497.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Trypsin.
DR Pfam: PF00089; trypsin.1.
DR PRINTS: PR00722; CHYMOTRYPSIN.

DR SMART; SM00020; TRYP-SPE; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 KW Hydrolyase; Serine protease.
 SQ SEQUENCE 262 AA; 27706 MW; DB4C1BFA6CA4EBBE CRC64;

Query Match 10.3%; Score 6; DB 5; Length 262;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 FPHQIS 24
 |||||
 DB 48 FPHQIS 53

RESULT 44
 ID 020667 PRELIMINARY; PRT; 264 AA.
 AC 020667;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE CODED FOR BY C. ELEGANS CDNA YK117H11.5.
 GN F52E1.1 OR POS-1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 CC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Faveilo A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Lalster N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson M., Weinstock L., Wilkinson-Spratt J., Wohlman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RT Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Le T.T., Rifkin L.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP MEDLINE=99054964; PubMed=9834181;
 RA Tabara H., Hill K.U., Mello C.C., Priess J.R., Kohara Y.;
 RT "pos-1 encodes a cytoplasmic zinc-finger protein essential for
 RT germline specification in C. elegans.";
 RL Development 126:1-11(1999).
 DR EMBL; U41109; AAB37036.1; -
 DR EMBL; AB006208; BAA33854.1; -
 DR InterPro; IPR000571; ZF-CCCH.
 DR Pfam; PF00642; ZF-CCCH; 2.
 DR SMART; SM00356; ZNF_C3H1; 2.
 KW Zinc.

SQ SEQUENCE 264 AA; 29830 MW; A9F52EC67B6B36C CRC64;
 Query Match 10.3%; Score 6; DB 5; Length 264;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KYKTVL 9
 |||||
 DB 141 KYKTVL 146

RESULT 45
 ID 09WZB9 PRELIMINARY; PRT; 268 AA.
 AC 09WZB9;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE CONSERVED HYPOTHETICAL PROTEIN.
 GN TM0651.
 OS Thermotoga maritima.
 CC Bacteria; Thermotogales; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MSB8 / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Ginn M.L., Dodson R.J.,
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of Thermotoga maritima.";
 RL Nature 399:323-329(1999).
 DR EMBL; AE001738; AAD35735.1; -
 DR TIGR; TM0651; -
 DR InterPro; IPR001454; Hydrolase.
 DR InterPro; IPR000150; Hypothet_cof.
 DR Pfam; PF00702; Hydrolase; 1.
 DR PROSITE; PS01228; COF_1; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 268 AA; 31186 MW; 02BB5753D5D46EED CRC64;

Query Match 10.3%; Score 6; DB 16; Length 268;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 IIEYIK 56
 |||||
 DB 91 IIEYIK 96

Search completed: November 5, 2002, 05:28:03
 Job time : 75 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 5, 2002, 03:10:33 ; Search time 1837 seconds

(without alignments)
660.718 Million cell updates/sec

Title: US-09-727-892A-99

Perfect score: 313

Sequence: 1 MERKRYKTVLVCDEIKGHFP.....YEVNLFTRKAYITIEYIKEI 58

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame-p2n.model -DEV=xlh
-O=/cgn2.1/USPTO/spool/US09727892/runat_01112002_185643_3860/app_query.fasta.1.199
-DB=GenEmbl -QFMT=fastap -SUFFIX=tran.rge -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi -LIT=45
-DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfo -HEAPSIZE=500 -MINTEN=0 -MAXLEN=2000000000
-USER=US09727892_ecgn.1.1182_erunat_01112002_185643_3860 -NCPU=6 -ICPU=3
-NO_XUPXT -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELext=7

Database :

GenEmbl:*
1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86.5	27.6	154821	8 AP003709	AP003709 Oryza sat
2	75.5	24.1	293431	2 PFMAL13P4	AL049181 Plasmodiu
3	75	24.0	21202	18 PFSC04088	AL1010273 Plasmodiu
4	75	24.0	61541	9 AC005232	AC005232 Homo sapi
5	75	24.0	135438	9 AC026444	AC026444 Homo sapi
6	75	24.0	149125	2 AC069401	AC069401 Homo sapi
7	75	24.0	224448	2 PFMAL4P4	AL035477 Plasmodiu
8	73.5	23.6	206462	2 AL645928	AL645928 Mus muscu
9	73.5	23.5	179966	9 AC026228	AC026228 Homo sapi
10	73.5	23.5	189760	9 AC022025	AC022025 Homo sapi
11	73.5	23.5	214828	2 AC099094	AC099094 Rattus no
12	72.5	23.2	160737	9 HS62R24	AL550003 Homo sapi
13	72	23.0	106591	2 AL592289	AL592289 Danio rer
14	72	23.0	178142	2 AC013546	AC013546 Homo sapi
15	72	23.0	180973	9 AC097462	AC097462 Homo sapi
16	72	23.0	183860	2 AC100723	AC100723 Mus muscu
17	72	23.0	185235	2 AL627076	AL627076 Mus muscu
18	72	23.0	236120	14 AF063866	AF063866 Melanoplu
19	71.5	22.8	169328	9 AP000472	AP000472 Homo sapi
20	71.5	22.8	180372	9 CNS01RG7	AL158058 Human chr
21	71.5	22.8	340000	9 AP001683	AP001683 Homo sapi
22	71	22.7	85132	9 UB2668	UB2668 Homo sapien
23	71	22.7	160536	9 AL355984	AL355984 Human DNA
24	71	22.7	188388	2 AC096875	AC096875 Pan trogl
25	70.5	22.5	122592	3 CEY60A3A	AL117207 Caenorhab
26	70.5	22.5	149886	2 AC106718	AC106718 Homo sapi
27	70.5	22.5	152537	9 AC055731	AC055731 Homo sapi
28	70.5	22.5	153549	2 AC022664	AC022664 Homo sapi
29	70.5	22.5	155909	2 AC097562	AC097562 Rattus no
30	70.5	22.5	161613	2 AC023947	AC023947 Homo sapi
31	70.5	22.5	163852	2 AC060795	AC060795 Homo sapi
32	70.5	22.5	188210	2 AC108719	AC108719 Homo sapi
33	70.5	22.5	191959	2 AC012217	AC012217 Homo sapi
34	70	22.4	4017	9 AB018262	AB018262 Homo sapi
35	70	22.4	28505	9 HSBAT75A3	AL132659 Human DNA
36	70	22.4	73193	2 AC102579	AC102579 Mus muscu
37	70	22.4	130299	8 AC019012	AC019012 Genomic S
38	70	22.4	159933	2 AC020923	AC020923 Homo sapi
39	70	22.4	163243	2 AC093003	AC093003 Homo sapi
40	70	22.4	166093	2 AC015462	AC015462 Homo sapi
41	70	22.4	174369	2 AC024267	AC024267 Homo sapi
42	70	22.4	200729	2 AC078783	AC078783 Homo sapi
43	70	22.4	215584	2 AL390779	AL390779 Homo sapi
44	70	22.4	312283	9 AF205588	AF205588 Homo sapi
45	69.5	22.2	10185	1 U32754	U32754 Haemophilus

ALIGNMENTS

RESULT 1
AP003709/c 154821 bp DNA linear PLN 17-OCT-2001
LOCUS Oryza sativa genomic DNA, chromosome 1, BAC clone:OSJNB0006H05,
DEFINITION complete sequence.
ACCESSION AP003709
VERSION AP003709.2 GI:16197552
KEYWORDS HTG.
SOURCE Oryza sativa (cultivar: Nipponbare) DNA, clone: OSJNB0006H05.

```

ORGANISM
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eleutherozoa; Oryzaceae; Oryza.
REFERENCE
AUTHORS
1 (bases 1 to 154821)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
ORIGIN
Oryza sativa nipponbare(GAS) genomic DNA, chromosome 1, BAC
clone:OSJNB0006H05
Published Only in Database (2001) In press
2 (bases 1 to 154821)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Direct Submission
Submitted (30-MAY-2001) Takuji Sasaki, National Institute of
Agricultural Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsakagi@nias.affrc.go.jp, URL:http://ygp.dna.affrc.go.jp/,
Tel:81-298-387441, Fax:81-298-387468)
On Oct 17, 2001 this sequence version replaced gi:14270109.
The orientation of the sequence is from M13rev to -21M13 of the BAC
clone.
FEATURES
source
1..154821
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="1"
/clone="OSJNB0006H05"
BASE COUNT
45329 a 33408 c 32630 g 43454 t
ORIGIN
Alignment Scores:
Pred. No.: 5.55 Length: 154821
Score: 86.50 Matches: 16
Percent Similarity: 60.87% Conservative: 12
Best Local Similarity: 34.78% Mismatches: 17
Query Match: 27.64% Indels: 1
DB: 8 Gaps: 1
US-09-727-892A-99 (1-58) x AP003709 (1-154821)
QY 10 LeuTYrCYsASPGLuILElYSGLyHSPhEPROHISGlnIleSeRMeCPheGUAsPLeU 29
:::||||| |:::: ||||| ::::
Db 110831 AATATATTGGTACTTTGAAGAACATATGAAATGCCCATATATTAATCAACGGTGCTACTATA 110772
QY 30 TYrAsPALALysVal---ValTySeerTYrTYrGUtYrASLauPheThiLYSLysfYr 48
||||| :::::| |::::| ::::
Db 110771 TATGAGAACCAAGTCGTACGTTGCTTACGTAACAAGTGTATATGACTTTCGCT 110712
QY 49 ALATyrILeILeGLUTyr 54
|||||:|||||:
Db 110711 GCATATATATGATTAT 110694
RESULT 2
PFMAL13P4
LOCUS
DEFINITION
Plasmodium falciparum chromosome 13 strain 3D7, *** SEQUENCING IN
PROGRESS ***, in unordered pieces.
ACCESSION
AL049181
VERSION
AL049181.4 GI:5731884
KEYWORDS
HTG: HTGS_PHASE1.
SOURCE
malaria parasite P. falciparum.
ORGANISM
Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE
Bowman,S., Churcher,C., Harris,B., Harris,D., Lawson,D., Quail,M.
and Barrell,B.
Direct Submission
Submitted (15-MAR-1999) P.falciparum Genome Sequencing Consortium,
The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA, UK
On Aug 12, 1999 this sequence version replaced gi:5513138.
For more information about this sequence or the Malaria Project,
see http://www.sanger.ac.uk/Projects/P.falciparum. IMPORTANT: This
COMMENT

```

	sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continues. The sequence may be contaminated with foreign sequence from E.coli, yeast, vector, page etc. Order of segments is not known; 800 n's separate segments. * NOTE: This is a 'working draft' sequence. * This record will be updated with the finished sequence * as soon as it is available and the accession number will be preserved.							
FEATURES	Location/Qualifiers							
source	1..293431 /organism= "Plasmodium falciparum" /strain= "3D7" /db_xref=taxon: 5833" /chromosome= "13"							
BASE COUNT	97622 a	23864 c	24769 g	94352 t	52824 others			
ORIGIN								
Alignment Scores:								
Pred. No.: Score:	283	Length:			293431			
Percent Similarity:	75.50	Matches:			21			
Best Local Similarity:	47.76%	Conservative:			11			
Query Match:	31.34%	Mismatches:			20			
DB:	24.12%	Indels:			15			
	2	Gaps:			3			
GS-09-727-892A-99 (1-58) x PFMAL13P4 (1-293431)								
OY 4 LysTyrLysThrValLeuLeuTYrCyAspScIleIeYsgLy-----	17							
Dd 107556 AAATTTAAATGTAAATATTATTCATGTGAAGTAACAAGGGGAAGCATGTATAACGCTTAA	107615							
OY 18 -----HisPhePCyHISGLNIlleSerMetPheGluspleutyraSpala 32								
Dd 107616 AAGAATATTATCAACACATATTATATAAAAATATTATATTATATAAAAATATTATATTAT	107675							
OY 33 LysValValTYrSerTYr-----TYrGIuTYr---AsnleuDheThrlYslYSlySTYr 48								
Dd 107676 AAAAATATATATATATATATATATATATATATATATATATATATATATATATATATATAT	107735							
OY 49 AlaTyrlilellegluTYrIIle 55								
Dd 107736 ATATATATATAAAAATATATATATATATATATATATATATATATATATATATATATAT	107756							
RESULT 3								
PFSC04088								
ID PFSC04088 standard; DNA; INV; 21202 BP.								
XX AC AL010273;								
XX SV AL010273.1								
SV 08-DEC-1997 (Rel. 53, Created)								
Dt 27-JUN-1998 (Rel. 56, Last updated, Version 2)								
Dc Plasmodium falciparum DNA *** SEQUENCING IN PROGRESS *** from contig 4-88								
XX HG.								
KM Plasmodium falciparum (malaria parasite P. falciparum)								
OS Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.								
OC Enkaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.								
XX [1]								
RN RP 1-21202								
RA Lawson D.;								
RT ;								
RL Submitted (25-JUN-1998) to the EMBL/GenBank/DBJ databases.								
RL P.falciparum Genome Sequencing Consortium, The Sanger Centre, Wellcome								
RL Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK								
CC IMPORTANT: This sequence is unfinished and does not necessarily								

CC represent the correct sequence. Work on the sequence is in progress
 CC and the release of this data is based on the understanding that the
 CC sequence may change as work continues. The sequence may be
 CC contaminated with foreign sequence from E.coli, yeast, vector,
 CC phage etc.

CC For more information about this sequence or the Malaria Project,
 CC see http://www.sanger.ac.uk/Projects/P_falciparum.

XX Key Location/Qualifiers
 FH source
 FT 1. 21202
 FT /chromosome="4"
 FT /db_xref="taxon:5833"
 FT /organism="Plasmodium falciparum"
 FT /clone="4-88"
 XX
 SO Sequence 21202 BP; 7427 A; 2248 C; 2061 G; 9466 T; 0 other;

Alignment Scores:
 Pred. No.: 22.4 Length: 21202
 Score: 75.00 Matches: 15
 Percent Similarity: 57.14% Conservative: 9
 Best Local Similarity: 35.71% Mismatches: 18
 Query Match: 23.96% Indels: 0
 DB: 18 Gaps: 0

US-09-727-892a-99 (1-58) x PFSC04088 (1-21202)

OY 11 Tyrcysaspgluilelysglyhispherohisglinilesermelphelglnaspleuayr 30
 ||| ||| ::||| ||| ||||| ||| |||||

Db 9507 TACATGATTTATTTGATAAAGATATTTTCGTCATATTTCTTTTTCATATCTGTTT 9566
 ||| ||| ::||| ||| ||||| ||| |||||

OY 31 Aspalalysvallyltyrsettyrtyrlytyrlyasnleuphehlysltyralalyr 50
 ||| ||| ::||| ||| ||||| ||| |||||

Db 9567 GATCATTTTGAATGTTTCATATTTATTCATATGTTTCATATTTTGACATATATATTTT 9626
 ||| ||| ::||| ||| ||||| ||| |||||

OY 51 lleile 52
 ::|

Db 9627 TTTTAA 9632
 ::|

RESULT 4
 AC005232/c 61541 bp DNA linear PRI 21-DEC-1999
 LOCUS
 DEFINITION Homo sapiens PAC clone RP5-1189D6 from 7p15.3-p14, complete
 sequence.
 ACCESSION AC005232
 VERSION AC005232.1 GI:4156156
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 1 (bases 1 to 61541)
 Mammalia; Eutheraia; Primates; Catarrhini; Homnidae; Homo.

AUTHORS
 Sulston,J.E. and Waterston,R.

TITLE
 Toward a complete human genome sequence

JOURNAL
 Genome Res. 8 (11), 1097-1108 (1998)

MEDLINE
 99063792

REFERENCE
 2 (bases 1 to 61541)
 Harmon,G., Kock,J. and Duckels,G.

AUTHORS
 The sequence of Homo sapiens PAC clone RP5-1189D6

JOURNAL
 Unpublished

REFERENCE
 3 (bases 1 to 61541)
 Waterston,R.H.

AUTHORS
 Direct Submission

JOURNAL
 Submitted (01-JUL-1998) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 4 (bases 1 to 61541)
 Waterston,R.H.
 Direct Submission
 Submitted (14-JAN-1999) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,

REFERENCE
 5 (bases 1 to 61541)
 Waterston,R.

AUTHORS
 Direct Submission

JOURNAL
 Submitted (19-AUG-1999) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

REFERENCE
 6 (bases 1 to 61541)
 Waterston,R.

AUTHORS
 Direct Submission

JOURNAL
 Submitted (21-DEC-1999) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Jan 14, 1999 this sequence version replaced gi:3309092.

COMMENT
 Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web Site: <http://genome.wustl.edu/gsc>
 Contact: saplens@watson.wustl.edu
 Summary Statistics
 Center project name: H_DJ1189D06

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and
 sequencing collaboration between the NHGRI Chromosome 7 Mapping
 Project (Eric D. Green, Director), John D. McPherson in the
 Department of Genetics (Washington University), and the Washington
 University Genome Sequencing Center. For additional information
 about the map position of this sequence, see
<http://www.nhgri.nih.gov/DIR/GTB/CHR7>, send
 mailto:egreen@nhgri.nih.gov, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

This clone was derived from human PAC library RPCR-5, prepared by
 Pieter de Jong and coworkers at the Roswell Park Cancer Institute
 (<http://pacpac.med.buffalo.edu>) using the method described by
 Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from
 one male donor.

The clone may be obtained either from Genome Systems, Inc.
 (<http://www.genomesystems.com>) or Research Genetics, Inc.
 (<http://www.resgen.com>); or from Pieter de Jong.

VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:
 The sequence of this clone is at base position 1 of RP5-1189D6;
 actual end is at 102833 of RP5-964C11.
 Location/Qualifiers

FEATURES

1. 61541
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7p15.3-p14"
 /clone="RP5-1189D6"
 /clone_1bp="RPCR-5"
 105..164
 repeat_region
 /rpt_family="MERL_type?"
 844..876
 repeat_region
 /rpt_family="POLX_A"
 877..904
 repeat_region
 /rpt_family="(TA)n"
 905..924
 repeat_region

```
repeat_region      15047. 15078
                    /rpt_family="MALR"
repeat_region      15081. 15403
                    /rpt_family="L2"
repeat_region      15404. 15715
                    /rpt_family="Alu"
repeat_region      15716. 15874
                    /rpt_family="L2"
repeat_region      15979. 16148
                    /rpt_family="MER90"
repeat_region      16455. 16550
                    /rpt_family="MIR"
repeat_region      17407. 17541
                    /rpt_family="Alu"
repeat_region      17656. 17736
                    /rpt_family="L2"
repeat_region      18413. 18454
                    /rpt_family="MIR"
repeat_region      18518. 18824
                    /rpt_family="Alu"
repeat_region      18883. 19175
                    /rpt_family="Alu"
repeat_region      19598. 19761
                    /rpt_family="L2"
```

Pred. No.:	66.5	Length:	61541
Score:	75.00	Matches:	17
Percent Similarity:	48.98%	Conservative:	7
Best Local Similarity:	34.69%	Mismatches:	25
Query Match:	23.96%	Indels:	0
DB:	9	Gaps:	0

9 LeuLeuTyrCysaspGluIleIleGlyHisPheProHisGlnIleSerMetPheGluasp 28

9399 CTGCTCAGATGTCACCTTACTTCAGGGGGGGCCCTTCCCTGACCACTATCTCAATAGTGCCA 9340

29 LeuTy rAspAla l y sValValTy rSe tTy rTy rGl uTy rAsnLeuPhe h rLy sLy sTy r 48
 ||| ||| ::||| ||| ||| ::||| |||
 Db 9339 CTGCTTACCCCTTCTCTCTATCCGTTATCCGCGTTGATTTTCTTAAACCTCTC 9280

QY	49	Ala	Tyr	Ile	Ile	Glu	Tyr	Ile	Glu	57
			:	:	:	:	:	:		
Db	9279	GCC	TAT	CTC	CTT	CC	TAC	GAT	TGA	9253

AC026444	AC026444	135438 bp	DNA	linear	PRI 01-MAY-2001
LOCUS	AC026444	135438 bp	DNA	linear	PRI 01-MAY-2001
DEFINITION	Homo sapiens chromosome 5 clone CTD-229958, complete sequence.				

ACCESSION	AC026444
VERSION	AC026444.5
	GI:13876493

KEYWORDS HTG.
SOURCE human.

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Homo

REFERENCE	
AUTHORS	Mammola; Euteria; Filinices; Catalinihi; Romlinuae; Romio.
DOE	1 (bases 1 to 135438)
Joint Genome Institute and Stanford Human genome Center	

INSTRUCTIONS	FOR CONTRIBUTORS	AND EDITORS	OF THE JOURNAL
DOE	DOE	DOE	DOE
TITLE	TITLE	TITLE	TITLE
Direct Submission	Direct Submission	Direct Submission	Direct Submission
Unpublished	Unpublished	Unpublished	Unpublished

2 (bases 1 to 135438)
DOE Joint Genome Institute.

TITLE Direct Submission
JOURNAL Submitted (22-MAR-2000) Production Sequencing Facility, DOE Joint

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94595
3 (bases 1 to 135438)

TITLE Direct Submission
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
ABSTRACT DOE Joint Genome Institute 2800 Mitchell
KEYWORDS Submitted (28-APR-2001)

REFERENCE 4 (bases 1 to 135438)
Cooknall, D. Submitted (20 Apr 2001) DOE Office
Drive, Walnut Creek, CA 94598, USA

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Apr 28, 2001 this sequence version replaced gi:13677029.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.hgsc.stanford.edu
Quality: phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.2.

FEATURES
source
1. 135438
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CTD-2299E8"
/clone="CTD-2299E8"

BASE COUNT 45448 a 24329 c 23699 g 41962 t
ORIGIN

Alignment Scores:
Pred. No.: 149 Length: 135438
Score: 75.00 Matches: 21
Percent Similarity: 51.79% Conservative: 8
Best Local Similarity: 37.50% Mismatches: 23
Query Match: 23.96% Indels: 5
DB: 9 Gaps: 1

US-09-727-892a-99 (1-58) x AC026444 (1-135438)

OY 7 ThrValLeuLeuTyrCysAspGluLeuLysGlyHisPheProHisGlnIle----- 23
||||| :|||:|||||:||||| :|||: |||
DB 26331 ACACATATATATATCTGAGGAGATATATATA-CACATATATATTTTATCTATATATA 26389
OY 24 ---SerMetPheGlnAspLeuTyrAspAlaLysValValTyrSerTyrTyrGluTyrasn 42
DB 26390 CTATATATATCTATATATCTCTATATAGTATGTTGATACATATATATATATACAT 26449
OY 43 LeuPheThrLysLysTyrAlaTyrIleLeuGluTyrIleLysGluIle 58
||||| :|||:|||||:||||| :|||: |||
DB 26450 ACATATCTCTATATATCTCTATATATCTCTATATATATATATATCTCTATA 26497

RESULT 6
AC069401/c 149125 bp DNA linear HTG-29-MAY-2000
LOCUS Homo sapiens chromosome 3 clone RP11-394L18 map 3p, WORKING DRAFT
DEFINITION
SEQUENCE, 53 unordered pieces.
AC069401
VERSION AC069401.1 GI:8101210
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 149125)
AUTHORS Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S.,
Huang,F., Jin,Y., Kang,N., Li,C., Li,G., Li,J., Li,L.,
Li,S., Li,T., Liu,Y., Liu,N., Li,B., Liu,Y., Li,W., Li,W., Li,Y.,
Luo,J., Miu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y.,
Tao,R., Wang,H., Wang,J., Wang,L., Wang,L., Wang,R.,
Wang,X., Wang,X., Wang,Y., Wu,Q., Xie,F., Xuan,Z., Xie,Y.,
Yao,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H.,
Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y.,
Zhang,Z., Zhu,B., Yu,J. and Yang,H.

TITLE
JOURNAL
REFERENCE
AUTHORS

COMMENT
TITLE
JOURNAL
Submitted (29-MAY-2000) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing,
100101, P.R.China
-----Genome Center
Center:Beijing Center
Center code:Beijing
Website:http://hgsc.igtp.ac.cn
http://www.genomics.org.cn
Contact:hgsc@igtp.ac.cn
-----Project Information
Center project name:1% project
Center clone name: RP11-394L18
-----Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator; ET 5% of reads
Chemistry: Dye-terminator Big Dye; 45% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 146353 bases at least Q40
Consensus quality: 156128 bases at least Q30
Consensus quality: 164031 bases at least Q20
Insert size: 96386; sum-of-continigs
Quality coverage: 4.73x in Q20 bases;sum-of-continigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 53 continigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the continigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1201: contig of 1201 bp in length
* 1202 1301: gap of unknown length
* 1302 2470: contig of 1169 bp in length
* 2471 2570: gap of unknown length
* 2571 3742: contig of 1172 bp in length
* 3743 3842: gap of unknown length
* 3843 5019: contig of 1177 bp in length
* 5020 5119: gap of unknown length
* 5120 6592: contig of 1473 bp in length
* 6593 7880: contig of 1188 bp in length
* 7881 7980: gap of unknown length
* 7981 9114: contig of 1134 bp in length
* 9115 9214: gap of unknown length
* 9215 11470: contig of 2256 bp in length
* 11471 12874: contig of 1304 bp in length
* 12875 12974: gap of unknown length
* 12975 14293: contig of 1319 bp in length
* 14294 14393: gap of unknown length
* 14394 15648: contig of 1255 bp in length
* 15649 17932: contig of 1284 bp in length
* 17933 18032: gap of unknown length
* 18033 19635: contig of 1603 bp in length
* 19636 19735: gap of unknown length
* 19736 21515: contig of 1780 bp in length
* 21516 22933: contig of 1318 bp in length
* 22934 23033: gap of unknown length
* 23034 24310: contig of 1277 bp in length
* 24311 24410: gap of unknown length
* 24411 26032: contig of 1622 bp in length
* 26033 27787: gap of unknown length
* 27788 27887: contig of 1655 bp in length
* 27887: gap of unknown length

27888	29140:	contig of 1253 bp in length
2911.1	29240:	gap of unknown length
2921.1	30401:	contig of 1161 bp in length
30402	30501:	gap of unknown length
30502	31793:	contig of 1292 bp in length
31794	31893:	gap of unknown length
31894	34390:	contig of 2497 bp in length
34391	34490:	gap of unknown length
34491	36556:	contig of 1866 bp in length
36337	36456:	gap of unknown length
36457	37895:	contig of 1439 bp in length
37896	37995:	gap of unknown length
37996	39366:	contig of 1371 bp in length
39367	39466:	gap of unknown length
39467	42862:	contig of 3396 bp in length
42863	42962:	gap of unknown length
42963	44735:	contig of 1773 bp in length
44736	44835:	gap of unknown length
44836	47409:	contig of 2574 bp in length
47410	47509:	gap of unknown length
47510	49561:	contig of 2052 bp in length
49562	49661:	gap of unknown length
49662	51627:	contig of 1966 bp in length
51628	51727:	gap of unknown length
51728	53869:	contig of 2142 bp in length
53870	53969:	gap of unknown length
53970	55304:	contig of 1335 bp in length
55305	55404:	gap of unknown length
55405	55694:	contig of 1290 bp in length
55695	56794:	gap of unknown length
56795	56267:	contig of 1473 bp in length
58268	58367:	gap of unknown length
58368	61293:	contig of 2926 bp in length
61294	61393:	gap of unknown length
61394	63767:	contig of 2374 bp in length
63768	63867:	gap of unknown length
63868	67089:	contig of 3222 bp in length
67090	67189:	gap of unknown length
67190	65999:	contig of 2810 bp in length
70000	70099:	gap of unknown length
70100	72976:	contig of 2877 bp in length
72977	73076:	gap of unknown length
73077	75848:	contig of 2772 bp in length
75849	75948:	gap of unknown length
75949	78561:	contig of 2613 bp in length
78562	78661:	gap of unknown length
78662	82088:	contig of 3427 bp in length
82089	82188:	gap of unknown length
82189	84724:	contig of 2536 bp in length
84725	84824:	gap of unknown length
84825	88484:	contig of 3660 bp in length
88485	88584:	gap of unknown length
88585	93065:	contig of 4481 bp in length
93066	93165:	gap of unknown length
93166	98240:	contig of 5075 bp in length
98241	98340:	gap of unknown length
98341	104246:	contig of 5906 bp in length
104247	104346:	gap of unknown length
104347	108994:	contig of 5548 bp in length
108995	109994:	gap of unknown length
109995	117646:	contig of 7652 bp in length
117647	117746:	gap of unknown length
117747	124817:	contig of 7071 bp in length
124818	124917:	gap of unknown length
124918	133446:	contig of 8529 bp in length
133447	133546:	gap of unknown length
133547	141346:	contig of 7800 bp in length
141347	141446:	gap of unknown length
141447	149125:	contig of 7679 bp in length

```
FEATURES      Location/Qualifiers
source        1. .149125
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
```

	/chromosome="3"	
	/map="3p"	
	/clone="RP11-394L18"	
	1..1201	
misc_feature	/note="assembly_name:Contig34"	
misc_feature	1302..2470	
misc_feature	/note="assembly_name:Contig35"	
misc_feature	2571..3742	
misc_feature	/note="assembly_name:Contig37"	
misc_feature	3843..5019	
misc_feature	/note="assembly_name:Contig43"	
misc_feature	5120..6592	
misc_feature	/note="assembly_name:Contig45"	
misc_feature	6693..7880	
misc_feature	/note="assembly_name:Contig47"	
misc_feature	7981..9114	
misc_feature	/note="assembly_name:Contig48"	
misc_feature	9215..11470	
misc_feature	/note="assembly_name:Contig51"	
misc_feature	11571..12874	
misc_feature	/note="assembly_name:Contig52"	
misc_feature	12975..14293	
misc_feature	/note="assembly_name:Contig53"	
misc_feature	14394..15648	
misc_feature	/note="assembly_name:Contig54"	

Alignment Scores:			
Pred. No.:	164	Length:	149125
Score:	75.00	Matches:	17
Percent Similarity:	48.98%	Conservative:	7
Best Local Similarity:	34.69%	Mismatches:	25
Query Match:	23.96%	Indels:	0
DB:	2	Gaps:	0

CS-09-727-892A-99 (1-58) x AC069401 (1-149125)

Qy	9	LeuTyrTCGyAspGlUlelYsgLyHISPherProHISglnIleSerMetRheglUasp	28
Dd	68952	CtGCCTCAGATGTCACTACTCATCGGGGGGCCCTTCCTGACACCATTCTCAAATAAGTGCCA	68893
Qy	29	LeuTYrASpAlaLYsValValTyreSeTYrTYrGLUTyrASnLeuPheThrLYsLYsTYr	48
Dd	68892	CtGCCTTAGACCCCTTTCCCTCATTCCTGCTTAATCCGTTATCTCGCTTGAATTTCTTAAGAAGCTCTC	68833
Qy	49	AlaTYrIlleIeGLUTyrIlleYsgLU	57
Dd	68832	GCCTATCTCTCTCTCCCTCAATTCAATGAA	68806

RESULT 7			
PFMALIP4/C			
DEFINITION	PFMALIP4	224448 bp	DNA linear HTG 11-AUG-1999
LOCUS	Plasmodium falciparum chromosome 4 strain 3D7, *** SEQUENCING IN		
PROGRESS ***, In unordered pieces.			
ACCESSION	AL035477		
VERSION	AL035477.5		
KEYWORDS	HTG; HTGS_PHASE1.		
SOURCE	malaria parasite P. falciparum.		
ORGANISM	Plasmodium falciparum		
REFERENCE	EduariVota, Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
AUTHORS	Bowman,S., Churchill,C., Harris,B., Harris,D., Lawson,D., Quail,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-FEB-1999) P.falciparum Genome Sequencing Consortium.		
COMMENT	The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK		
COMMENT	On Aug 12, 1999 this sequence version replaced gi:5531400.		
COMMENT	For more information about this sequence or the Malaria Project,		
COMMENT	see http://www.sanger.ac.uk/Projects/P_falciparum.IMPORTANT: This		
COMMENT	sequence is unfinished and does not necessarily represent the		
COMMENT	correct sequence. Work on the sequence is in progress and the		
COMMENT	release of this data is based on the understanding that the		
COMMENT	sequence may change as work continues. The sequence may be		

contaminated with foreign sequence from E.coli, yeast, vector, phage etc.

Order of segments is not known: 800 n's separate segments.

* NOTE: This is a 'working draft' sequence.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

FEATURES
source 1..224448
/organism="Plasmodium falciparum"
/strain="3D7"
/db_xref="taxon:5833"
/chromosome="4"

BASE COUNT 80792 a 23258 c 21957 g 75240 t 23201 others
ORIGIN

Alignment Scores:

Pred. No.:	250	Length:	224448
Score:	75.00	Matches:	15
Percent Similarity:	57.14%	Conservative:	9
Best Local Similarity:	35.71%	Mismatches:	18
Query Match:	23.96%	Indels:	0
DB:	2	Gaps:	0

US-09-727-892A-99 (1-58) x PFMA14P4 (1-224448)

QY 11 TTYCTGASAGLIIleLysGlyHisPheProHisGlnIleSerMetPheGluAspLeuTyr 30

DB 169677 TACATGAGATTATGTAAGAAAGATTTCGTACATATATTTCTTTTTCATATCTGTTT 169618

QY 31 ASpalalysValIleValtyrSerTyrTyrGluTyrAsnLeuPheThrLysLysTyrAlaTyr 50

DB 169617 GATCATTTTGCATATGTTTTCATATATATCATATATGTTTCATATATTTGATCATATATTTT 169558

QY 51 llelle 52

DB 169557 TTTTTA 169552

RESULT 8

AL645928/c 206462 bp DNA linear HTG 15-JAN-2002

LOCUS Mus musculus chromosome 11 clone RP23-185J21, *** SEQUENCING IN

PROGRESS ***, in ordered pieces.

ACCESSION AL645928 8 GI:18181747

VERSION HTG: HTGS_PHA5E2; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS house mouse.

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (sites)

Tracey/A.

Direct Submission

Submitted (14-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk

On Jan 16, 2002 this sequence version replaced gi:18072571.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquerry@sanger.ac.uk

----- Project Information

Center project name: BM185021

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Sequencing vector: plasmid; 108752; 100% of reads

Consensus quality: 206377 bases at least Q40

Consensus quality: 206404 bases at least Q20

Insert size: 206462; sum-of-contigs

Quality coverage: 14.60x in Q20 bases; sum-of-contigs quality

coverage: 14.71x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

FEATURES
source 1..206462
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-185J21"
/clone_1b="RPC1-23"
1..206462
/note="assembly_fragment:02312
clone_end:17
vector_side:left"

misc_feature

BASE COUNT 61773 a 40600 c 39969 g 64120 t
ORIGIN

Alignment Scores:

Pred. No.:	309	Length:	206462
Score:	74.00	Matches:	19
Percent Similarity:	51.85%	Conservative:	9
Best Local Similarity:	35.19%	Mismatches:	20
Query Match:	23.64%	Indels:	6
DB:	2	Gaps:	3

US-09-727-892A-99 (1-58) x AL645928 (1-206462)

QY 8 ValIleuLeuTyr---CysaspIuIle-----LysGlyHisPheProHisGln 22

DB 72519 TTGTTACTATATGCTGTTCTTCTGTTATTTATTTACCAAGGTCACCATTTCTCAAGG 72460

QY 23 lIleSerMetPheGluAspLeuTyrAspAla---LysValIleValtyrSerTyrGluTyr 41

DB 72459 TTAATATGTTATCTATACATACAGATCTTTTAAGATTTAAGAGCGCATTTGTGATATAC 72400

QY 42 AsnLeuPheThrLysLysTyrAlaTyrIlelleGluTyrIle 55

DB 72399 ACAATATTATATACATATATATATATTAACCTACATA 72358

RESULT 9

AC026228/c 179966 bp DNA linear PRI 24-OCT-2001

LOCUS Homo sapiens chromosome 10 clone RP11-95G14, complete sequence.

ACCESSION AC026228 4 GI:16356838

VERSION HTG.

KEYWORDS human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 179966)

Smith,D.R.

Genome Therapeutics Corporation Sequencing Center: Human Genome

Sequence Data

Unpublished

Submitted (21-MAR-2000) Genome Therapeutics Corporation, 100 Beaver

Street, Waltham, MA 02453, USA

3 (bases 1 to 179966)

Smith,D.R.

Direct Submission

Submitted (24-OCT-2001) Genome Therapeutics Corporation, 100 Beaver

Street, Waltham, MA 02453, USA

On Oct 24, 2001 this sequence version replaced gi:13786258.

----- Location/Qualifiers

1..179966

/organism="Homo sapiens"

```

/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-95G14"
/clone_11b="RPC1-11"
BASE COUNT      60748 a 33098 c 31772 g 54347 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:      312      Length:      179966
Score:          73.50     Matches:      19
Percent Similarity: 60.00%  Conservative: 11
Best Local Similarity: 38.00%  Mismatches:  13
Query Match:     23.48%     Indels:      7
DB:              9        Gaps:         3

US-09-727-892A-99 (1-58) x AC026228 (1-179966)

QY      2  GUAATGlyTyrLysThValLeuLeuTyrCysAspGluIleGlyHisPheProHis 21
DB      12696 CAAAGAAAGATATTAATGACTGACACACATGC---AAGCTAAAGGAAACTTTGATCAG 12640
QY      22  GlnIleSerMetPheGluAspLeuTyrAspAlaLysValValTyrSerTyrTyrGluTyr 41
DB      12639 -----TCACATTTCCACACAGTTA-----CTTGATATCATCTTTTATGAAATAC 12598
QY      42  AsnLeuPheThrLysLysTyrAlaTyrIle 51
DB      12597 TTAGTATACACGACGACCTTGATGATACATC 12568

RESULT 10
AC022025/c      189760 bp      DNA      linear      PRI 15-DEC-2001
LOCUS      AC022025      Homo sapiens chromosome 10 clone RP11-489D8, complete sequence.
DEFINITION      AC022025
ACCESSION      AC022025
VERSION      AC022025.6  GI:17861011
KEYWORDS      HTG.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 189760)
AUTHORS      Smith,D.R.
TITLE      Genome Therapeutics Corporation Sequencing Center: Human Genome
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 189760)
AUTHORS      Smith,D.R.
TITLE      Direct Submission
JOURNAL      Submitted (25-JAN-2000) Genome Therapeutics Corporation, 100 Beaver
REFERENCE      3 (bases 1 to 189760)
AUTHORS      Smith,D.R.
TITLE      Direct Submission
JOURNAL      Submitted (15-DEC-2001) Genome Therapeutics Corporation, 100 Beaver
COMMENT      Street, Waltham, MA 02453, USA
FEATURES
     source
     1..189760
         location/Qualifiers
         1. .189760
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /chromosome="10"
             /clone="RP11-489D8"
             /clone_11b="RPC1-11"
BASE COUNT      61011 a 36277 c 33262 g 59210 t
ORIGIN

Alignment Scores:
Pred. No.:      329      Length:      189760
Score:          73.50     Matches:      19
Percent Similarity: 60.00%  Conservative: 11
Best Local Similarity: 38.00%  Mismatches:  13
Query Match:     23.48%     Indels:      7
DB:              9        Gaps:         3

```

```

US-09-727-892A-99 (1-58) x AC022025 (1-189760)
QY      2  GUAATGlyTyrLysThValLeuLeuTyrCysAspGluIleGlyHisPheProHis 21
DB      15395 CAAAGAAAGATATTAATGACTGACACACATGC---AAGCTAAAGGAAACTTTGATCAG 153939
QY      22  GlnIleSerMetPheGluAspLeuTyrAspAlaLysValValTyrSerTyrTyrGluTyr 41
DB      15398 -----TCACATTTCCACACAGTTA-----CTTGATATCATCTTTTATGAAATAC 153897
QY      42  AsnLeuPheThrLysLysTyrAlaTyrIle 51
DB      153896 TTAGTATACACGACGACCTTGATGATACATC 153867

RESULT 11
AC090904/c      214828 bp      DNA      linear      HTG 21-DEC-2001
LOCUS      AC090904      Rattus norvegicus chromosome R14 clone CH230-191E19, WORKING DRAFT
DEFINITION      SEQUENCE, 13 unordered pieces.
ACCESSION      AC090904
VERSION      AC090904.3  GI:17975727
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Norway rat.
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE      1 (bases 1 to 214828)
AUTHORS      Muzny,D.M., Adams,C., Adio-Oduola,B., All-osman,F.R., Allen,C.,
Aisbrook,S.L., Amaralunge,H.C., Aye,R., Banks,T., Barbria,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,R.,
Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dethorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douhwaite,K.J., Draper,H.,
Dugan-Rocha,S., Dublin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Franz,P., Gabisl,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guvera,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodson,A., Hognes,M., Holloway,C.,
Hollins,B., Homsli,F., Howard,S., Huber,J., Huliy,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvan,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Meshawari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokweto,S.,
Ogun,M., Okunolu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Prims,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojurokan,I., Rolfe,M.,
Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shooshitari,N.,
Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabar,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Washington,S., Williams,G., Williamson,A., Wleczky,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G., and Gibbs,R.
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 214828)
AUTHORS      Worley,K.C.
TITLE      Direct Submission

```

JOURNAL

COMMENT

Submitted (09-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Dec 21, 2001 this sequence version replaced gi:17062411.

Center: Baylor College of Medicine
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GGN1
Center clone name: CH230-191E19

Summary Statistics

Assembly program: Phrap; version 0.990329first call to findPhrapList

Consensus quality: 208083 bases at least Q40
Consensus quality: 210269 bases at least Q30
Consensus quality: 211939 bases at least Q20
Estimated insert size: 210041; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 6.8x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently consists of 13 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 72737: contig of 72737 bp in length
72738 72837: gap of unknown length
72838 11203: contig of 38366 bp in length
11204 11103: gap of unknown length
11104 134804: contig of 23501 bp in length
134805 134904: gap of unknown length
134905 157908: contig of 23004 bp in length
157909 158009: gap of unknown length
158010 174036: contig of 16028 bp in length
174037 174136: gap of unknown length
174137 189112: contig of 14976 bp in length
189113 189212: gap of unknown length
189213 201169: contig of 11957 bp in length
201170 201269: gap of unknown length
201270 207176: contig of 5907 bp in length
207177 207276: gap of unknown length
207277 209081: contig of 1805 bp in length
209082 209181: gap of unknown length
209182 210432: contig of 1251 bp in length
210433 210532: gap of unknown length
210533 211825: contig of 1293 bp in length
211826 211925: gap of unknown length
211926 213165: contig of 1240 bp in length
213166 213265: gap of unknown length
213266 214828: contig of 1563 bp in length.

FEATURES
Source
1. 214828
Location/Qualifiers

/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/chromosome="Rf4"
/clone="CH230-191E19"

BASE COUNT 59515 a 47914 c 45504 g 60652 t 1243 others
ORIGIN

Alignment Scores:

Pred. No.: 373 Length: 214828
Score: 73.50 Matches: 18
Percent Similarity: 47.54% Conservative: 11
Best Local Similarity: 29.51% Mismatches: 27
Query Match: 23.48% Indels: 5
DB: 2 Gaps: 1

US-09-727-892a-99 (1-58) x AC099094 (1-214828)

Qy 2 GlnArglystYrlystThValLeuLeuYrlysp-----GlnIlelys 16

Db 205054 CAGCGCAAGGAGAGACCTCTTATTCGATTTTAAATACGATATCACTCA 204995

Qy 17 GlyHisPheProHisGlnIleSerMetPheGlnAspLeuYrlyspAlaValValTyr 36

Db 204994 TATCATATACACACACGATATATATATATATATATATATATATATATAT 204935

Qy 37 SerTyrYrlystYrlyspLeuPheThrlystYrlystYrlyleuIleGluTyrlylys 56

Db 204934 ATATATATATATATATATATATATATATATATATATATATATATATATG 204875

Qy 57 Glu 57

Db 204874 GAA 204872

RESULT 12

HS62F24

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission
Submitted (09-NOV-2000) GBF, Dept. of Genome Analysis, Mascheroder Weg 1, D-38124 Braunschweig, Germany, E-mail: info.genome@gbf.de
On Feb 1, 2002 this sequence version replaced gi:11138114.
All annotations in this database entry are developed by computational tools. It is therefore not explicitly noted in the feature lines that evidence is not experimental.
Mapping was performed at The Sanger Centre
(cf. <http://www.sanger.ac.uk/HGP/chr9>)
Mapping information is available via
<http://webace.sanger.ac.uk/cgi-bin/display?db=acedb9&grep=62F24>
----- Genome Center
Center: GBF, Braunschweig
Center code: GBF
Web site: <http://genome.gbf.de/>
Contact: info.genome@gbf.de
----- Project Information
Center project name: ba62F24
Center clone name: ba62F24
----- Summary Statistics
Sequencing vector: ##;
Chemistry: Dye-terminator-BigDye; 46% of reads
Chemistry: Dye-terminator-ambersham; 45% of reads
Chemistry: Dye-primer-ambersham; 9% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 0 bases at least Q40
Consensus quality: 0 bases at least Q30
Consensus quality: 0 bases at least Q20
Estimated insert size: ##; agarose-fp estimation
Estimated insert size: 160737; sum-of-contigs estimation

PROGRAMS AND PARAMETERS USED FOR ANNOTATION:
++++++
+ Analysis and annotation were performed with the automatic +


```

/note="TTTG repeat"
repeat_region 11280..11533

Alignment Scores:
Pred. No.: 374 Length: 160737
Score: 72.50 Matches: 16
Percent Similarity: 52.00% Conservative: 10
Best Local Similarity: 32.00% Mismatches: 13
Query Match: 23.16% Indels: 11
DB: 9 Gaps: 2

US-09-727-892a-99 (1-58) x HS62F24 (1-160737)
Oy 3 ArglySTyRlySThValleuLeuYrCysaspGuilelySGlyHIsPheProHIsGln 22
Db 19441 AGAAGGCAATAG-----CTTACCTGCATGAGCTTTATGTCATTTCCCATATAT 19491
Oy 23 lIeSer-----MetPheGluaspLeuTyraSPAlaLysVal 34
Db 19492 CTTAAGTGTACAAATAATCTGCTCCTATTTGTACAAAGCTAGTGTGATTTAAATAT 19551
Oy 35 ValTySerTyTyRglUtyrAsnLeuPhe 44
Db 19552 AAATTATCTATTTTAAATATTTATT 19581

RESULT 13
AL592289/c 106591 bp DNA linear HTG 02-FEB-2002
LOCUS Dario rerio clone XX-1CSE, *** SEQUENCING IN PROGRESS ***, 1n
DEFINITION ordered pieces.
ACCESSION AL592289
VERSION AL592289.10 GI:184496251
KEYWORDS HTG; HTGS_PHASE2; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE zebrafish.
ORGANISM Dario rerio.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (sites)
REFERENCE
AUTHORS Hammond, S.
TITLE Direct Submission
JOURNAL Submitted (01-FEB-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquyer@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 4, 2002 this sequence version replaced gi:18476717.
COMMENT
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquyer@sanger.ac.uk
----- Project Information
Center project name: BAC1CSE
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 104880 bases at least Q40
Consensus quality: 105297 bases at least Q30
Consensus quality: 105767 bases at least Q20
Insert size: 106591; sum-of-contigs
Insert size: 103493; 12.0% error; agarose-fp
Quality coverage: 13.33x in Q20 bases; sum-of-contigs quality
coverage: 13.72x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
Location/Qualifiers
1..106591
/organism="Dario rerio"
/db_xref="taxon:7955"
/clone="XX-1CSE"
/clone_1fb="Zeb_BAC"
FEATURES
SOURCE

```

```

misc_feature 1..106591
/note="assembly_fragment:03752
clone_end:SP6
vector_side:left"
BASE COUNT 33226 a 19033 c 19756 g 34576 t
ORIGIN
-----
Alignment Scores:
Pred. No.: 285 Length: 106591
Score: 72.00 Matches: 18
Percent Similarity: 43.08% Conservative: 10
Best Local Similarity: 27.69% Mismatches: 19
Query Match: 23.00% Indels: 18
DB: 2 Gaps: 3

US-09-727-892a-99 (1-58) x AL592289 (1-106591)
Oy 3 ArglySTyRlySThValleuLeuYrCysaspGuilelySGlyHIsPhe----- 19
Db 32773 AAAAGTACACCGGTTATTCTACTGAC-----TTACGGGGTCATTTTCGTGATGAA 32723
Oy 20 -----ProHIsGlnIleSerMetPheGluaspLeuTyraSP 31
Db 32722 CATCTGTTGTACCCCTTTTACCAAGCATTAAGGTTGTCCCTCATACCTAC--- 32666
Oy 32 AlalysValValTySerTyTyRglUtyrAsnLeuPheThyRlySTyTAlaTyrlle 51
Db 32665 -----ACATACGTTCAATTATCTTACACAGTACGATGACGAGAAATACACTTAAT 32615
Oy 52 lIeGluTyrlleLys 56
Db 32614 TCACGTACTTACGC 32600

RESULT 14
AC013546/c 178142 bp DNA linear HTG 31-OCT-2001
LOCUS Homo sapiens chromosome 8 clone RP11-431D12 map 8, *** SEQUENCING
DEFINITION IN PROGRESS ***, 1 ordered pieces.
ACCESSION AC013546
VERSION AC013546.8 GI:16041437
KEYWORDS HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 178142)
REFERENCE
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 11, 2001 this sequence version replaced gi:15963680.
COMMENT
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research

```


1:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1960.DAT.*
2:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1961.DAT.*
3:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1962.DAT.*
4:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1963.DAT.*
5:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1964.DAT.*
6:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1965.DAT.*
7:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1966.DAT.*
8:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1967.DAT.*
9:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1968.DAT.*
10:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1969.DAT.*
11:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1990.DAT.*
12:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1991.DAT.*
13:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1992.DAT.*
14:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1993.DAT.*
15:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1994.DAT.*
16:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1995.DAT.*
17:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1996.DAT.*
18:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1997.DAT.*
19:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1998.DAT.*
20:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA1999.DAT.*
21:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA2000.DAT.*
22:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA2001.DAT.*
23:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA2002.DAT.*
24:	/SIDS1/gcgdata/genseq/genseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	% Match	Query Length	ID	Description
1	313	100.0	177	AAA69042	Bacteriophage 44AH
2	313	100.0	16668	AAA69013	Bacteriophage 44AH
3	283	90.4	165	AAA69045	Bacteriophage 44AH
4	71	22.7	15577	AAV35616	SHOX gene prelinin
5	71	22.7	32367	AAV35620	Human SHOX (short
6	70	22.4	2142	AAAF72840	Secreted protein 9
7	70	22.4	2404	AAK94441	Human full-length
8	70	22.4	2868	AAK94441	Human full-length
9	70	22.4	3696	AAAF72827	Secreted protein 9
10	68	21.7	2985	AAK94463	Human full-length
11	66	21.1	1683	AAH51955	Mycobacterium tube
12	66	21.1	4403765	AAI99683	Mycobacterium tube
13	66	21.1	4411529	AAI99682	Mycobacterium tube
14	65.5	20.9	3514	ABLI4760	Drosophila melanog
15	65	20.8	1234	AAAC77698	Human cancer assoc
16	65	20.8	1237	AAAC7463	Arabidopsis thalia
17	65	20.8	1240	AAAC33460	Arabidopsis thalia
18	65	20.8	113515	ABLI34174	Human immune syste
19	64.5	20.6	397	AAAS39309	Novel human diagno
20	63	20.1	337	AAI89179	Human polynucleoti
21	63	20.1	1329	AAAS5502	Streptococcus pneu
22	63	20.1	1329	AAAS5567	Streptococcus pneu
23	63	20.1	1948	AAAS2768	Corn putative cata
24	63	20.1	3722	ABLI4044	Drosophila melanog
25	63	20.1	3831	AAV52424	Streptococcus pneu
26	63	20.1	4927	ABLI4038	Drosophila melanog
27	63	20.1	6132	ABLI32863	Human immune syste
28	62.5	20.0	4993	ABLI4103	Human immune syste
29	62.5	20.0	4997	AAAS8823	cDNA encoding a hu
30	62.5	20.0	5641	ABLI3396	Human immune syste
31	62.5	20.0	134499	AAAF32286	BAC containing rep
32	62	19.8	636	AAH73216	Human cervical can
33	62	19.8	1512	AAH297062	Human secreted pro
34	62	19.8	3639	AAAS64109	Human prostate CDN
35	62	19.8	3639	AAH93873	p788P full length
36	62	19.8	3848	AAH27277	Prostate-tumour de
37	62	19.8	4646	AAH49924	Human calcium chan
38	62	19.8	5668	AAH27262	Prostate-tumour de
39	62	19.8	5668	AAAS64108	Human prostate CDN
40	62	19.8	5668	AAH93872	p788P full length
41	61.5	19.6	3540	ABLI9100	Drosophila melanog
42	61.5	19.6	4651	AAH46688	DNA encoding an al
43	61	19.5	1321	AAH37128	Human PRO1781 (UNQ
44	61	19.5	1321	AAAS6110	Human DNA encoding
45	61	19.5	1321	AAFS4458	Probe #50 used in

RESULT 1

AAA69042

ID AAA69042 standard; DNA; 177 BP.

AAA69042;

27-OCT-2000 (first entry)

Bacteriophage 44AHND nucleotide sequence 44AHNDORF025.

Bacteriophage; antimicrobial; genome; identification; antibacterial;

bacterial growth inhibition; bacterial infection; ds.

Bacteriophage 44AHND.

WO200032825-A2.

08-JUN-2000.

ALIGNMENTS

```
XX 03-DEC-1999; 99WO-IB02040.
XX
XX 03-DEC-1998; 98US-0110992.
XX
XX 03-JUN-1999; 99US-0326144.
XX
XX 28-SEP-1999; 99US-0407804.
XX
XX 30-SEP-1999; 99US-0157218.
XX
XX 01-DEC-1999; 99US-0168777.
XX
XX 02-DEC-1999; 99US-0454252.
XX
XX (PHAG-) PHAGETECH INC.
XX
XX Pelletier J, Gros P, Dubow M;
XX
XX WPI; 2000-412361/35.
XX
XX P-PSDB; AAB16557.
XX
XX
XX Identifying a bacteriophage coding region for treating bacterial
XX infections comprises identifying a nucleic acid encoding a product that
XX inhibits bacteria when a bacteriophage infects a bacterium
XX
XX Example 9; Page 278; 456pp; English.
XX
XX The present invention describes a method for identifying a bacteriophage
XX coding region encoding a product active on an essential bacterial
XX target. The method comprises identifying a nucleic acid sequence encoding
XX a gene product that provides a bacteria-inhibiting function when an
XX uncharacterised bacteriophage infects a pathogenic bacterium. The
XX compound active on a target of a bacteriophage inhibitor protein in a
XX bacteria is used to treat or prevent a bacterial infection in an animal.
XX AAB68243 to AAB69442 and AAB16523 to AAB16954 represent bacteriophage
XX nucleotide and protein sequences which are used in the exemplification of
XX the present invention.
XX
XX SQ Sequence 177 BP; 72 A; 22 C; 26 G; 57 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 2,69e-34 Length: 177
XX Score: 313.00 Matches: 58
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 21 Gaps: 0
XX
XX US-09-727-892A-99 (1-58) x AAA69042 (1-177)
XX
XX QY 1 MetGLuARGLySTyRlYsThrValleuLeuTYrCYsAspGLuileYsGLYHisPhePro 20
XX |||||||
XX DB 1 ATGGAAACGTAATACAAACACGGTATTATTATTCGATGAGATTAAAGCATTTTCCA 60
XX
XX QY 21 HisGLuIleSerMetPheGLuAspLeuTYrAspAlaLYsValValTYrSerTYrTYrGLu 40
XX |||||||
XX DB 61 CATCAATCTCAATGTTTGAAGATTATATGACGCTAAAGTGTATTCATATTATGAA 120
XX
XX QY 41 TYrAsnLeuPheThrLYsLYsTYrAlaTYrIleleGLuTYrIleLYsGLuile 58
XX |||||||
XX DB 121 TATTAACCTGTCTACTAAAAAATACGCGTATATCATAGAAATACATTAAAGAGATA 174
XX
XX RESULT 2
XX AAA69013/c
XX ID AAA69013 standard; DNA: 16668 BP.
XX
XX AC AAA69013;
XX
XX DT 27-OCT-2000 (first entry)
XX
XX DE Bacteriophage 44AHD complete genome sequence.
XX
XX DE Bacteriophage 44AHD nucleotide sequence 44AHDORF029.
XX
XX DE Bacteriophage; antimicrobial; genome; identification; antibacterial;
XX
XX KW bacterial growth inhibition; bacterial infection; ds.
XX
XX OS Bacteriophage 44AHD.
XX
XX
```

```
XX W0200032825-A2.
XX
XX 08-JUN-2000.
XX
XX 03-DEC-1999; 99WO-IB02040.
XX
XX
XX 03-DEC-1998; 98US-0110992.
XX
XX 03-JUN-1999; 99US-0326144.
XX
XX 28-SEP-1999; 99US-0407804.
XX
XX 30-SEP-1999; 99US-0157218.
XX
XX 01-DEC-1999; 99US-0168777.
XX
XX 02-DEC-1999; 99US-0454252.
XX
XX (PHAG-) PHAGETECH INC.
XX
XX Pelletier J, Gros P, Dubow M;
XX
XX WPI; 2000-412361/35.
XX
XX
XX Identifying a bacteriophage coding region for treating bacterial
XX infections comprises identifying a nucleic acid encoding a product that
XX inhibits bacteria when a bacteriophage infects a bacterium
XX
XX Example 9; Page 266-269; 456pp; English.
XX
XX The present invention describes a method for identifying a bacteriophage
XX coding region encoding a product active on an essential bacterial
XX target. The method comprises identifying a nucleic acid sequence encoding
XX a gene product that provides a bacteria-inhibiting function when an
XX uncharacterised bacteriophage infects a pathogenic bacterium. The
XX compound active on a target of a bacteriophage inhibitor protein in a
XX bacteria is used to treat or prevent a bacterial infection in an animal.
XX AAB68243 to AAB69442 and AAB16523 to AAB16954 represent bacteriophage
XX nucleotide and protein sequences which are used in the exemplification of
XX the present invention.
XX
XX SQ Sequence 16668 BP; 6095 A; 2338 C; 2608 G; 5627 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 6,6e-32 Length: 16668
XX Score: 313.00 Matches: 58
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 21 Gaps: 0
XX
XX US-09-727-892A-99 (1-58) x AAA69013 (1-16668)
XX
XX QY 1 MetGLuARGLySTyRlYsThrValleuLeuTYrCYsAspGLuileYsGLYHisPhePro 20
XX |||||||
XX DB 15175 ATGGAAACGTAATACAAACACGGTATTATTATTCGATGAGATTAAAGCATTTTCCA 15116
XX
XX QY 21 HisGLuIleSerMetPheGLuAspLeuTYrAspAlaLYsValValTYrSerTYrTYrGLu 40
XX |||||||
XX DB 15115 CATCAATCTCAATGTTTGAAGATTATATGACGCTAAAGTGTATTCATATTATGAA 15056
XX
XX QY 41 TYrAsnLeuPheThrLYsLYsTYrAlaTYrIleleGLuTYrIleLYsGLuile 58
XX |||||||
XX DB 15055 TATTAACCTGTCTACTAAAAAATACGCGTATATCATAGAAATACATTAAAGAGATA 15002
XX
XX RESULT 3
XX AAA69045
XX ID AAA69045 standard; DNA: 165 BP.
XX
XX AC AAA69045;
XX
XX DT 27-OCT-2000 (first entry)
XX
XX DE Bacteriophage 44AHD complete genome sequence.
XX
XX DE Bacteriophage 44AHD nucleotide sequence 44AHDORF029.
XX
XX DE Bacteriophage; antimicrobial; genome; identification; antibacterial;
XX
XX KW bacterial growth inhibition; bacterial infection; ds.
XX
XX
```

OS Bacteriophage 44AHJD.
 XX WO200032825-A2.
 PN
 XX
 XX 08-JUN-2000.
 PD
 XX
 XX 03-DEC-1999; 99WO-IB02040.
 PF
 XX
 XX 03-DEC-1998; 98US-0110992.
 PR 03-JUN-1999; 99US-0326144.
 PR 28-SEP-1999; 99US-0407804.
 PR 30-SEP-1999; 99US-0157218.
 PR 01-DEC-1999; 99US-0168777.
 PR 02-DEC-1999; 99US-0454252.
 XX
 PA (PHAG-) PHAGETECH INC.
 XX
 PI Pelletier J, Gros P, Dubow M;
 XX
 DR WPI, 2000-412361/35.
 DR P-PSDB; AAB16560.
 XX
 XX
 PT Identifying a bacteriophage coding region for treating bacterial
 PT infections comprises identifying a nucleic acid encoding a product that
 PT inhibits bacteria when a bacteriophage infects a bacterium
 XX
 PS Example 9; Page 278; 456pp; English.
 XX
 CC The present invention describes a method for identifying a bacteriophage
 CC coding region encoding a product active on an essential bacterial
 CC target. The method comprises identifying a nucleic acid sequence encoding
 CC a gene product that provides a bacteria-inhibiting function when an
 CC uncharacterized bacteriophage infects a pathogenic bacterium. The
 CC compound active on a target of a bacteriophage inhibitor protein in a
 CC bacteria is used to treat or prevent a bacterial infection in an animal.
 CC AAB68243 to AAB69442 and AAB16523 to AAB16954 represent bacteriophage
 CC nucleotide and protein sequences which are used in the exemplification of
 CC the present invention.
 XX
 SQ Sequence 165 BP; 63 A; 21 C; 25 G; 56 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 3.01e-30 Length: 165
 Score: 283.00 Matches: 52
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 90.42% Indels: 0
 DB: 21 Gaps: 0
 US-09-727-892A-99 (1-58) x AAB69045 (1-165)
 OY 1 MetGluAqgLyStyLysThrValLeuLeuTYrCysAspGluIlleLysGlyHisPhePro 20
 DB 9 AAGGACGTAATTAACAAAACGGTATTATATTCGAGATTAAGGACATTTTCCA 68
 OY 21 HisGlnIleSerMetPheGluAspLeuTYrAspAlaLysValValTyrSerTyrGlu 40
 DB 69 CATCAATCTCAATGTTTGACATTTATATGACGCTAAAGTTGATTTCAATATATGAA 128
 OY 41 TyrAsnLeuPheThrLysLysTyrAlaTyrIleIle 52
 DB 129 TATAACCTGTTCACTAAAAAATACCGTATATCATA 164
 RESULT 4
 ID AAV35616 standard; DNA; 15577 BP.
 AC AAV35616;
 XX
 XX 07-SEP-1998 (first entry)
 XX
 DE SHOX gene preliminary nucleotide sequence (H0X93).
 XX

KW Homeobox domain; human growth gene; growth regulation; growth defect;
 KW Turner's syndrome; short stature homeobox containing gene; short stature;
 KW SHOX; bone disease; osteoporosis; calcium regulation; H0X93; ss.
 XX
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH exon 1498..1807
 FT /tag= a
 FT /note= "part of exon I (G310)"
 FT 3844..4068
 FT /tag= b
 FT /note= "PBT92 region (first part)"
 FT 4326..4437
 FT /tag= c
 FT /note= "PBT92 region (second part)"
 FT 4543..4619
 FT /tag= d
 FT /note= "PBT92 region (third part)"
 FT 5305..5512
 FT /tag= e
 FT /note= "part of exon II (ET93)"
 FT 11620..11729
 FT /tag= f
 FT /note= "part of exon IV (G108)"
 XX
 PN W09814568-A1.
 XX
 PD 09-APR-1998.
 XX
 XX 29-SEP-1997; 97WO-EP05355.
 XX
 PR 16-JAN-1997; 97EP-0100583.
 PR 01-OCT-1996; 96US-0027633.
 XX
 XX (RAPP/) RAPPOLD-HOERBRAND G.
 PA
 PI Rao E, Rappold-hoerbrand G;
 XX
 DR WPI, 1998-271719/24.
 XX
 PT New human growth genes - used to develop products for the diagnosis
 PT and treatment of human growth defects such as short stature, e.g.
 PT Turner's syndrome
 XX
 PS Disclosure: Pages 37-45; 84pp; English.
 XX
 CC This is a preliminary nucleotide sequence of the SHOX gene. The gene
 CC region corresponding to short stature has been identified as a region of
 CC approximately 500 kb in the PAR1 region of the X and Y chromosomes.
 CC Three genes in this region have been identified as candidates for the
 CC short stature gene. These genes were designated SHOX (also referred to
 CC as SHOX93 or H0X93), PBT92 and SHOT (SHOX-like homeobox gene on
 CC chromosome three). The SHOX gene has two separate splicing sites
 CC resulting in two variations SHOXa and SHOXb. The specification provides
 CC sequences of SHOX (short stature homeobox-containing) genes SHOX ET92,
 CC SHOXa, SHOXb, SHOT and exons of the SHOX genes as shown in AAV35610 to
 CC AAV35621 and protein sequences of the human growth protein transcription
 CC factor SHOXa, SHOXb and SHOT as shown AAW60573 to AAW60575. The novel
 CC genes are responsible for human growth. Defects in the genes can cause
 CC short stature, e.g. Turner's syndrome. Defects in the genes can cause
 CC develop agents for the treatment of short stature or other human growth
 CC disorders. The products can also be used for providing a mitogenic effect
 CC on cells, e.g. for the treatment of bone diseases such as osteoporosis
 CC and diseases involved with disturbance in the bone calcium regulation.
 XX
 SQ Sequence 15577 BP; 3692 A; 3875 C; 4196 G; 3629 T; 185 other;
 XX
 Alignment Scores:
 Pred. No.: 55.9 Length: 15577
 Score: 71.00 Matches: 20
 Percent Similarity: 52.63% Conservative: 10
 Best Local Similarity: 35.09% Mismatches: 19

Query Match	22.68%	Indels:	8
DB:	19	Gaps:	3
DB: 05-09-727-892A-99 (1-58) x AAV35616 (1-15577)			
OY 5	TYRLysThrValleuLeuTYrCysAspGluIle-----LysGlyHisPheProHis 21		
DB 7063	TMTAAATATATACAAACATTTGTGATATATATATATAATATATATAAAACATATATACAT 7122		
OY 22	GluIleSerMetPheGluAspLeuTYrAspAlaLysValTYrSerTYrTYrGluIuYr 41		
DB 7123	---ATTAATAATATATATATAACATATATACATATATTAAGAAATATAT-----ATA 7167		
OY 42	AsnLeuPheThrLysLysTYrAlaTYrIleIleGluTYrIleLysGluIle 58		
DB 7168	AACATATATACATATATAATATACATATATAAACATATATATACATATAAATA 7218		
RESULT 5			
AAV35620			
ID AAV35620	standard; DNA; 32367 BP.		
AC AAV35620;			
XX 07-SEP-1998	(first entry)		
XX Human SHOX	(short stature homeobox containing gene) gene sequence.		
XX Homeobox domain; human growth gene; growth regulation; growth defect;			
XX Turner's syndrome; short stature homeobox containing gene; SHOXa;			
XX SHOX; bone disease; osteoporosis; calcium regulation; short stature;			
XX transcription factor A; ss.			
XX Homo sapiens.			
OS WO9814568-A1.			
PN 09-APR-1998.			
PD 29-SEP-1997;	97WO-BE03535.		
PF 16-JAN-1997;	97EP-0100583.		
PR 01-OCT-1996;	96US-0027633.		
XX (RAPD/) RAPPOLD-HOERRAND G.			
PA Rao E, Rapppold-hoerbrand G;			
XX WPI: 1998-271719/24.			
PI New human growth genes - used to develop products for the diagnosis			
DR and treatment of human growth defects such as short stature, e.g.			
PT Turner's syndrome			
XX Claim 19; Pages 51-67; 84pp: English.			
XX This is the human SHOX gene sequence containing the PARI region. The			
CC gene region corresponding to short stature has been identified as a			
CC region of approximately 500 kb in the PARI region of the x and y			
CC chromosomes. Three genes in this region have been identified as			
CC candidates for the short stature gene. These genes were designated SHOX			
CC (also referred to as SHOX93 or HOS93), per92 and SHOT (SHOX-like homeobox			
CC gene on chromosome three). The SHOX gene has two separate splicing sites			
CC resulting in two variations SHOXa and SHOXb. The specification provides			
CC sequences of SHOX (short stature homeobox-containing) genes SHOX ET92,			
CC SHOXa, SHOXb, SHOT and exons of the SHOX genes as shown in AAV35616 to			
CC AAV35621 and protein sequences of the human growth protein transcription			
CC factor SHOXa, SHOXb and SHOT as shown AAW60573 to AAW60575. The novel			
CC genes are responsible for human growth. Defects in the genes can cause			
CC short stature, e.g. Turner's syndrome. The products can be used to			
CC develop agents for the treatment of short stature or other human growth			
CC disorders. The products can also be used for providing a mitogenic effect			
CC on cells, e.g. for the treatment of bone diseases such as osteoporosis			
CC and diseases involved with disturbance in the bone calcium regulation.			

XX	Seq	Sequence 32367 BP; 7627 A; 8130 C; 8564 G; 8043 T; 3 other;
XX	Alignment Scores:	
XX	Pred. No.:	135
XX	Score:	71.00
XX	Percent Similarity:	52.63%
XX	Best Local Similarity:	35.09%
XX	Query Match:	22.68%
XX	DB:	19 Gaps: 3
XX	OS-09-727-892A-99 (1-58) x AAV35620 (1-32367)	
OY	5 TyrIysThrValLeuLeuTyrCysAspGluIle-----LysGIlyHisPheProHis 21	
DB	6950 TATAAATATTAACCAACAATATTGTATATATATATATATATAAACATATATATCAT	
OY	22 GluIleSerMetPheGluAspLeuTyrAspAlaAlaValValTyrSerTyrTyrgluTyr 41	
DB	7010 ---ATAAAAAATATATATATTAACATATATACATATATTAAGAAATATAT-----ATA 7055	
OY	42 AsnLeuPheThrLysTyrAlaTyrIleIleGluTyrIleLysGluIle 58	
DB	7055 AACATATATACATATATAATATACATATATTAACATATATATACATATAATA 7105	
RESULT 6		
AAF72840/C		
ID	AAF72840 standard; DNA; 2142 BP.	
XX	AAF72840;	
XX	24-APR-2001 (first entry)	
DE	Secreted protein gene #42.	
KW	Secreted protein; human; autoimmune; hyperproliferation;	
KW	cardiovascular; cerebrovascular; infection; food; ds.	
XX	Homo sapiens.	
PN	MO200107459-A1.	
PD	01-FEB-2001.	
PF	20-JUL-2000; 2000WO-US19735.	
PR	23-JUL-1999; 99US-0145220.	
PA	(HUMA-) HUMAN GENOME SCI INC.	
PI	Rosen CA, Ruben SM, Ebner R, Duan RD, Ni J, Soppet DR, Moore PA,	
PI	Shi Y, Lafleur DW, Olsen HS, Birse CE, Komatsoulis GA;	
DR	WPI; 2001-123261/13.	
PT	New isolated nucleic acid encoding 29 secreted proteins, for	
PT	diagnosing, preventing and treating e.g. autoimmune,	
PT	hyperproliferative, cardiovascular, and ocular diseases or disorders	
PT	and microorganism infections -	
PS	Claim 1; Page 526-527; 601pp; English.	
CC	The present invention relates to 29 human secreted proteins. The	
CC	invention is used to prevent autoimmune diseases e.g. rheumatoid	
CC	arthritis, hyperproliferative disorders e.g. neoplasms of the	
CC	breast or liver, cardiovascular disorders e.g. cardiac arrest,	
CC	cerebrovascular disorders e.g. cerebral ischemia, angiogenesis,	
CC	nervous system disorders e.g. Alzheimer's disease, infections	
CC	caused by bacteria, viruses and fungi and ocular disorders e.g.	
CC	corneal infection. Also used in food preparations.	
SO	Sequence 2142 BP; 639 A; 346 C; 441 G; 715 T; 1 other;	

Alignment Scores:

Pred. No.:	6.92	Length:	214.2
Score:	70.00	Matches:	13
Percent Similarity:	53.85%	Conservative:	15
Best Local Similarity:	22.00%	Mismatches:	20
Query Match:	22.36%	Indels:	4
DB:	22	Gaps:	1

US-09-727-892A-99 (1-58) x AAF72840 (1-2142)

Qy 7 ThrValLeuLeuTyrCysAspGluIleLysGlyHisPheProHisGlnIleSerMetPhe 26

Alignment Scores:

Pred. No.:	7.96	Length:	240
Score:	70.00	Matches:	13
Percent Similarity:	53.85%	Conservative:	15
Best local Similarity:	25.00%	Mismatches:	20
Query Match:	22.36%	Indels:	4
DB:	22	Gaps:	1

US-09-727-892A-99 (1-58) x AAK94911 (1-2404)

QY 7 ThrValLeuLeuTyrCysaspIuTeluSglYhisPheProHISGlnIleSerMetPhe 26

Alignment Scores:	
Pred. No.:	9, 86
Score:	70.00
Percent Similarity:	53.85%
Best Local Similarity:	25.00%
Query Match:	22,368
DB:	22
Length:	2868
Matches:	15
Conservative:	13
Mismatches:	20
Indels:	4
Gaps:	1

US-09-727-892A-99 (1-58) x AAK94441 (1-2868)

QY 7 ThrValLeuLeuYrCysAspRgLIleLysGlyHisPheProHISglInIleSerMetPhe 26
 ||||| : : : ||| | | | | | |
Db 1528 ACACtACTGTTCCTCCCTTAACTCTGTGGTGCACTCCACCACATTAACAAGAcTTTC 1469

QY 27 GluAspLeuYrAspAlaLysValValTySerTyTrGluTyLysAsnLeuPheThrLys 46
 ::: ||||| : : : | | | | | |
Db 1468 CAAcACATTAATTCATTGA-----TATTTGGCATCTGCATTAATTTACTGCA 1421

QY 47 LysTyralaTyrllelleGluTyrlleLysGluIle 58
 ::: : : : : : | | | | |
Db 1420 AGAAATCTGCTGTGTGGAGCTTTAAAGAACAAATGTA 1385

Score:	70.00	Matches:	13
Percent Similarity:	53.05%	Conservative:	15
Best Local Similarity:	25.00%	Mismatches:	20
Query Match:	22.36%	Indels:	4
DB:	22	Gaps:	1

cy	7	Thallaleu	leu	tyr	cys	asn	glu	ile	lys	gln	his	phe	pro	his	gln	ile	ser	thr	phe	26	
Db	2292	ACAC	TAC	GTG	TTT	CA	CTT	TAC	CTG	CTG	CTG	CA	CTC	ACC	CCC	ATC	CA	TAC	CA	CTC	2233
cy	27	glu	asp	leu	tyr	asp	ala	lys	val	val	tyr	ser	tyr	tyr	glu	tyr	asn	phe	thr	lys	46
Db	2232	CAC	ACA	CTA	TAT	CA	TCA	ATT	A	GA	-----	TAT	TTT	GCA	TCT	GCA	TAT	ATT	A	CTCA	2185
cy	47	lys	tyr	ala	tyr	ile	ile	leu	gln	tyr	ile	lys	gln	ile	58						
Db	2184	AGA	AA	AT	CT	CT	CT	GT	G	TG	G	CA	CTT	TA	AG	CA	AA	CT	A	CTA	2149

Pred. No.: 19.4 Length: 2985
 Score: 68.00 Matches: 13
 Percent Similarity: 53.85% Conservative: 15
 Best Local Similarity: 25.00% Mismatches: 20
 Query Match: 21.73% Indels: 4
 DB: 22 Gaps: 1
 US-09-727-892a-99 (1-58) x AAH51959 (1-2985)
 QY 7 ThrValLeuLeuTyrCysAspGluIleLysGlyHisPheProHisGlnIleSerMetPhe 26
 DB 1645 ACACGTACTGTTCTTACCTCTTCTGTCGTCACCTCCCATCATCACTCAGTCTTC 1586
 QY 27 GluAspLeuTyrSepAlaLysValValTyrSerTyrTyrGluTyrAsnLeuPheThrLys 46
 DB 1585 CACACATATATCAATTAAGAA-----TATTTTGATCTGCATTAATTAAGTCA 1538
 QY 47 LysTyrAlaTyrIleIleGluTyrIleLysGluIle 58
 DB 1537 AGAAATCTGCTGTTATGAGCTTAAGAGAAATGTA 1502
 RESULT 11
 AAH51959 standard; DNA: 1683 BP.
 AC AAH51959;
 XX
 DT 04-SEP-2001 (first entry)
 XX
 DE Mycobacterium tuberculosis potential drug target gene SEQ ID 13.
 XX
 KM Drug target; growth; organism viability; characterisation; ds.
 XX
 OS Mycobacterium tuberculosis.
 PN WO200135317-A1.
 PD 17-MAY-2001.
 PF 13-NOV-2000; 2000MO-US31152.
 PR 12-NOV-1999; 9905-0165086.
 PR 12-NOV-1999; 9905-0165124.
 PR 01-FEB-2000; 2000US-0179531.
 XX
 PA (REGC) UNIV CALIFORNIA.
 PI Eisenberg D, Rotstein SH, Marcotte EW;
 DR WPI: 2001-329193/34.
 DR P-PSDB; AAG81108.
 XX
 PT Identifying nucleotide or polypeptide sequence for use as drug target,
 PT involves providing algorithm that analyzes a functional relationship
 PT between nucleotide or polypeptide sequences, and comparing the
 PT sequences -
 PS
 PS Disclosure; Page 56; 207pp; English.
 XX
 CC This invention relates to a method for identifying a nucleotide or
 CC polypeptide sequence that may be a drug target, or essential for growth
 CC or viability of an organism. Polynucleotide sequences AAH51947 - AAH52092
 CC represent DNA encoding proteins AAG81096 - AAG81241. Mycobacterium
 CC tuberculosis proteins which are potential drug targets. The DNA and
 CC protein sequences are used to illustrate the method of the invention. The
 CC method involves providing an unknown nucleotide or polypeptide sequences,
 CC and comparing it to a number of sequences along with at least one
 CC algorithm capable of analyzing a functional relationship between
 CC nucleotide and polypeptide sequences. The method is useful for
 CC characterizing the function of nucleic acids and polypeptides that may be
 CC useful as a target for a drug or essential for the growth or viability of
 CC an organism.
 CC
 CC
 CC

SQ Sequence 1683 BP; 239 A; 518 C; 568 G; 358 T; 0 other;
 Alignment Scores:
 Pred. No.: 18.1 Length: 1683
 Score: 66.00 Matches: 18
 Percent Similarity: 47.27% Conservative: 8
 Best Local Similarity: 32.73% Mismatches: 19
 Query Match: 21.09% Indels: 10
 DB: 22 Gaps: 2
 US-09-727-892a-99 (1-58) x AAH51959 (1-1683)
 QY 7 ThrValLeuLeuTyrCysAspGluIle-----LysGlyHisPheProHisGlnIleSer 24
 DB 337 ACGGTGCCGCTCTATGGCGCTGAGTTATTGAGCCAAAGCAAGTTTCCGTACAAATCAAGC 396
 QY 25 MetPheGluAsp-----LeuTyrAspAlaLysValValTyr 36
 DB 397 TGGATCGAAGACGACAGACACCGCAGCTGCGCTACGAGACAGATCGCGGTG 456
 QY 37 SerTyrTyrGluTyrAsnLeuPheThrLysTyrAlaTyrIle 51
 DB 457 CGCTATATGAGATCCGCTGCTGATCTGATCATCATCATCTG 501
 RESULT 12
 AA199683 standard; DNA: 4403765 BP.
 ID AA199683;
 XX
 AC AA199683;
 XX
 DT 15-JAN-2002 (first entry)
 XX
 DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
 XX
 KM Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
 KM variation; epidemiology; patient treatment; epidemic monitoring; ds.
 XX
 OS Mycobacterium tuberculosis.
 PN US6294328-B1.
 PD 25-SEP-2001.
 PF 24-JUN-1998; 9805-0103840.
 PR 24-JUN-1998; 9805-0103840.
 XX
 PA (GENO-) INST GENOMIC RES.
 PI Fleischmann RD, White OR, Fraser CM, Venter JC;
 DR WPI: 2001-647261/74.
 DR
 XX
 PT Evaluating strain variation of Mycobacterium tuberculosis, comprises
 PT determining the nucleotide sequence of the strain at positions in the
 PT genome corresponding to positions where M. tuberculosis strains CDC
 PT 1551 and H37Rv differ -
 PS
 PS Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
 XX
 CC The invention relates to evaluating strain variation within and between
 CC different populations of the tuberculosis bacterial pathogen,
 CC Mycobacterium tuberculosis or related Mycobacterium by determining the
 CC nucleotide sequence of the first strain at positions in the complete
 CC sequence of the genome that correspond to positions that differ in the
 CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AA199683) and
 CC H37Rv (AA199682). The method is useful for evaluating strain variation of
 CC M. tuberculosis and has valuable application in the fields of
 CC tuberculosis genetics, epidemiology, patient treatment and epidemic
 CC monitoring.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from USPRO
 CC at seqdata.uspto.gov/sequence.html?docID=6294328B1.
 CC
 CC
 CC

XX SQ Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;
Alignment Scores:
Pred. No.: 2.4e+05 Length: 4403765
Score: 66.00 Matches: 18
Percent Similarity: 47.27% Conservative: 8
Best Local Similarity: 32.73% Mismatches: 19
Query Match: 21.09% Indels: 10
Gaps: 2
DB: 22
US-09-727-892a-99 (1-58) x AA199683 (1-4403765)
QY 7 ThrValLeuLeuTyrCysAspGluIle-----LysGlyHisPheProHisGlnIleSer 24
DB 55972 ACGGTGCGCGCTCTATGCGCGCTGAGTTATGAGCCAGCAAGTTCCGTCAAAATCAAGC 56031
QY 25 MetPheGluAsp-----LeuTyrAspAlaLysValValTyr 36
DB 56032 TGGATCGAAACCGACAGCAAGCCAGCCAGCTGCGCTACGAGGACAGATCGCGGTG 56091
QY 37 SerTyrTyrGluTyrAsnLeuPheThrLysLysTyrAlaTyrIle 51
DB 56092 CGCTATATGAGATATCGGTGCTGACTGAGATCTATCAGTACCTG 56136
RESULT 13
AA199682
ID AA199682 standard; DNA; 4411529 BP.
XX AC AA199682;
XX DF 15-JAN-2002 (first entry)
XX DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
XX KM Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
XX variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX OS Mycobacterium tuberculosis.
XX PN US6294328-B1.
XX PD 25-SEP-2001.
XX PF 24-JUN-1998; 98US-0103840.
XX PR 24-JUN-1998; 98US-0103840.
XX PA (GENO-) INST GENOMIC RES.
XX PI Fleischmann RD, White OR, Fraser CM, Venter JC;
XX WPI: 2001-647261/74.
XX DR
XX PT Evaluating strain variation of Mycobacterium tuberculosis, comprises
XX PT determining the nucleotide sequence of the strain at positions in the
XX PT genome corresponding to positions where M. tuberculosis strains CDC
XX PT 1551 and H37Rv differ -
XX PS Claim 3; SEQ ID NO 1; 3bp + Sequence Listing; English.
XX
XX The invention relates to evaluating strain variation within and between
XX CC different populations of the tuberculosis bacterial pathogen,
XX CC Mycobacterium tuberculosis or related Mycobacterium by determining the
XX CC nucleotide sequence of the first strain at positions in the complete
XX CC sequence of the genome that correspond to positions that differ in the
XX CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AA199683) and
XX CC H37Rv (AA199682). The method is useful for evaluating strain variation of
XX CC M. tuberculosis and has valuable application in the fields of
XX CC tuberculosis genetics, epidemiology, patient treatment and epidemic
XX CC monitoring.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from USPTO

CC at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX SQ Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;
Alignment Scores:
Pred. No.: 2.41e+05 Length: 4411529
Score: 66.00 Matches: 18
Percent Similarity: 47.27% Conservative: 8
Best Local Similarity: 32.73% Mismatches: 19
Query Match: 21.09% Indels: 10
Gaps: 2
DB: 22
US-09-727-892a-99 (1-58) x AA199682 (1-4411529)
QY 7 ThrValLeuLeuTyrCysAspGluIle-----LysGlyHisPheProHisGlnIleSer 24
DB 56030 ACGGTGCGCGCTCTATGCGCGCTGAGTTATGAGCCAGCAAGTTCCGTCAAAATCAAGC 56089
QY 25 MetPheGluAsp-----LeuTyrAspAlaLysValValTyr 36
DB 56090 TGGATCGAAACCGACAGCAAGCCAGCCAGCTGCGCTACGAGGACAGATCGCGGTG 56149
QY 37 SerTyrTyrGluTyrAsnLeuPheThrLysLysTyrAlaTyrIle 51
DB 56150 CGCTATATGAGATATCGGTGCTGACTGAGATCTATCAGTACCTG 56194
RESULT 14
ABL14760
ID ABL14760 standard; cDNA; 3514 BP.
XX AC ABL14760;
XX AC 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 38762.
XX KM Drosophila; developmental biology; cell signalling; insecticide;
XX KM pharmaceutical; gene; ss.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001MO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EM;
XX WPI: 2001-656860/75.
XX DR P-PSDB; ABB70657.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions -
XX PS Claim 1; SEQ ID NO 38762; 21bp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-AB130511), expressed DNA
XX CC sequences (ABL101840-ABL16175) and the encoded proteins
XX CC (ABBS7737-ABB72072).
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 3514 BP; 921 A; 817 C; 838 G; 938 T; 0 other;
Alignment Scores:
Pred. No.: 51.7 Length: 3514
Score: 65.50 Matches: 16
Percent Similarity: 41.43% Conservative: 13
Best Local Similarity: 22.86% Mismatches: 26
Query Match: 20.93% Indels: 15
DB: 23 Gaps: 1
US-09-727-892a-99 (1-58) x ABL14760 (1-3514)
QY 3 ArgysrYrlysrThrValleuLeuTyrcysAspGluileLysGlyHisPheProHisGln 22
DB 3064 CGCAATrTAATAAAGGCTCCGCGTCACAGCAAACTAAACCTGGATATTTTGAAGGGCGA 3123
QY 23 lleserMetPheGluAspLeuTyrcysPala----- 32
DB 3124 CTATCCAAATGTGCAAGCTTGTGGATGTCATTTCCAAAGGGGAAATACTATTAAACT 3183
QY 33 -----LysValValTyrcSerTyrcGluTyrcAsnLeuPheThrLysLys 47
DB 3184 CGCAACGAAAGAGCGCAATGTGAACATACATCTTAAACCACTTCTTGTCAAATA 3243
QY 48 TyralaTyrllelleGluTyrcLysGlu 57
DB 3244 AAGACTTGGTAGTGAATATACATCAGAA 3273
RESULT 15
AAC77698
ID AAC77698 standard; cDNA; 1234 BP.
XX
AC AAC77698;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human cancer associated gene sequence SEQ ID NO:92.
XX
KW Human; cancer associated gene; cancer antigen; detection; cancer;
diagnosis; cytostatic; proliferative; vulnerrary; immunomodulator;
antidiabetic; antidiabetic; antidiabetic; antidiabetic; antiviral;
antidiabetic; antidiabetic; antidiabetic; antidiabetic; cardiant;
dermatological; neuroprotective; thrombolytic; coagulant; noctropic;
KW vasotropic; antipruritic; antidiabetic; gene therapy; inflammation;
immune disorder; hematopoietic cell disorder; autoimmune disorder;
allergic reaction; graft versus host disease; organ rejection;
KW hemostatic; thrombolytic; cardiovascular disorder; infection;
neurological disease; drug screening; ss.
XX
OS Homo sapiens.
XX
PN WO20005350-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05882.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587533/55.
XX
DR P-PSDB; AAB43489.
XX
PT Novel isolated nucleic acids comprising sequences encoding peptides
useful for treating or diagnosing e.g. cancer -
XX
PS Claim 1; Page 683; 2352pp; English.

CC AAC77607 to AAC78448 encode the human cancer associated proteins given
CC in AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnerrary; immunomodulator;
CC antidiabetic; antidiabetic; antidiabetic; antidiabetic;
CC antidiabetic; antidiabetic; antidiabetic; antidiabetic; antiviral;
CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC noctropic; vasotropic; antipruritic and antidiabetic;
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention.

XX S0 Sequence 1234 BP; 410 A; 239 C; 254 G; 330 T; 1 other;

Alignment Scores:
Pred. No.: 17 Length: 1234
Score: 65.00 Matches: 22
Percent Similarity: 54.84% Conservative: 12
Best Local Similarity: 35.48% Mismatches: 14
Query Match: 20.77% Indels: 14
DB: 21 Gaps: 4

US-09-727-892a-99 (1-58) x AAC77698 (1-1234)

QY 4 LysTyrcLysThrValleuLeuTyrcysAspGluileLysGlyHisPhe----- 19
DB 623 AAATTCAGGCTGTGATCTCTATACAGAGAAAGCCCACTTCTTCCCTTAATAC 682
QY 20 -----ProHisGlnIleSerMetPheGluAspLeuTyrcAsnLeuPheThrLysLys 36
DB 683 GTTCCAGCTGTGACAGCATGATATCTATACGATAT---GATGCGAGCGTGGCA 739
QY 37 SerTyrcTyrcGluTyrcAsnLeuPheThrLysLysTyrcAlaTyrcLysGlu 55
DB 740 AATTACCAAGATATCAATCTG-----GCCACATCATCTACTACTCTCTG 784
QY 56 LysGlu 57
DB 785 AAGGAG 790

Search completed: November 5, 2002, 05:15:29
Job time : 7380 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 5, 2002, 03:07:43 ; Search time 46 Seconds

(without alignments)
309.712 Million cell updates/sec

Title: US-09-727-892A-99

Perfect score: 313

Sequence: 1 MERRKYTVLLYCEIKHFP.....YEVNLTFRKYAYIIYEIKEL 58

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=x1h
-O=/cgn2_1/USPRO.spool/US09727892/runat_01112002_185641_3847/app-query.fasta.1.199
-DB=Issued_Patents.NA -OPM=fastap -SUFFIX=tran.rn1 -MINMATCH=0.1 -LOOPEXT=0
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09727892_CGN1_1_13_01una_01112002_185641_3847 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAPP -LARGEOUTERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents.NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCPUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66	21.1	4403765	4	US-09-103-840A-2
2	66	21.1	4411529	4	US-09-103-840A-1
3	62	19.8	3848	4	US-09-112-096-28
4	62	19.8	5668	4	US-09-112-096-14
5	61	19.5	1984	1	US-07-885-970A-25
6	61	19.5	1985	1	US-08-298-687A-25
7	61	19.5	1985	1	US-08-298-829-25
8	60.5	19.3	2133	3	US-08-947-965-1
9	59.5	19.0	476	4	US-08-642-274D-25
10	59.5	19.0	476	4	US-08-952-014C-25
11	59.5	19.0	3507	1	US-08-315-468-3
12	59	18.8	630	2	US-08-743-637B-173

c	13	59	18.8	630	3	US-08-526-840B-173	Sequence 173, App
	14	59	18.8	2875	3	US-09-255-893-1	Sequence 1, Appl1
	15	59	18.8	4946	3	US-08-930-996A-1	Sequence 1, Appl1
	16	58.5	18.7	438	1	US-08-688-609-5	Sequence 5, Appl1
	17	58.5	18.7	438	1	US-09-002-832-5	Sequence 5, Appl1
	18	58.5	18.7	444	1	US-08-688-609-2	Sequence 2, Appl1
	19	58.5	18.7	444	3	US-09-002-832-2	Sequence 2, Appl1
	20	58.5	18.7	587	3	US-08-688-609-4	Sequence 4, Appl1
	21	58.5	18.7	587	3	US-09-002-832-4	Sequence 4, Appl1
	22	58.5	18.7	830	3	US-08-688-609-1	Sequence 1, Appl1
	23	58.5	18.7	830	3	US-09-002-832-1	Sequence 1, Appl1
	24	58	18.5	5394	3	US-08-688-376-1	Sequence 1, Appl1
	25	57	18.2	709	1	US-08-413-135-2	Sequence 2, Appl1
	26	57	18.2	709	4	US-08-971-395-2	Sequence 2, Appl1
	27	57	18.2	3182	4	US-08-971-395-1	Sequence 1, Appl1
	28	57	18.2	3183	1	US-08-413-135-1	Sequence 1, Appl1
	29	57	18.2	3705	5	PCT-US96-03940-7	Sequence 7, Appl1
	30	57	18.2	5648	5	PCT-US96-03940-1	Sequence 1, Appl1
	31	56.5	18.1	1847	4	US-09-381-849-5	Sequence 5, Appl1
	32	56.5	18.1	2531	3	US-08-714-918-60	Sequence 60, Appl1
	33	56.5	18.1	2531	4	US-09-265-315-60	Sequence 60, Appl1
	34	56.5	18.1	2531	4	US-09-265-315-60	Sequence 60, Appl1
	35	56.5	18.1	2531	4	US-09-265-315-60	Sequence 60, Appl1
	36	56.5	18.1	3160	4	US-08-936-165A-255	Sequence 255, App
	37	56.5	18.1	3659	1	US-08-220-151-72	Sequence 72, Appl
	38	56.5	18.1	3659	1	US-08-303-124-12	Sequence 12, Appl
	39	56.5	18.1	3659	1	US-08-413-118-72	Sequence 72, Appl
	40	56.5	18.1	3659	1	US-08-224-657-48	Sequence 48, Appl
	41	56.5	18.1	3659	2	US-08-184-009-77	Sequence 77, Appl
	42	56.5	18.1	3659	2	US-08-417-210A-67	Sequence 67, Appl
	43	56.5	18.1	3659	2	US-08-480-697B-12	Sequence 12, Appl
	44	56.5	18.1	3659	2	US-08-458-356-77	Sequence 77, Appl
	45	56.5	18.1	3659	3	US-08-473-446-72	Sequence 72, Appl

ALIGNMENTS

RESULT 1
US-09-103-840A-2
Sequence 2, Application US/09103840A
Patent No. 6294328

GENERAL INFORMATION:
APPLICANT: FRASER, Claire M.
APPLICANT: WHITE, Owen R.

APPLICANT: FRASER, Claire M.
VENTER, John C.

TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TUBERCULOSIS

FILE REFERENCE: 24366-20007.00
CURRENT FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2
LENGTH: 4403765

TYPE: DNA
ORGANISM: Mycobacterium tuberculosis

FEATURE:
OTHER INFORMATION: CDC 1551

OTHER INFORMATION: "n" bases at various positions throughout the sequence
and is derived by analysis of the total score distribution.

US-09-103-840A-2

Alignment Scores:
Pred. No.: 3.56e+04
Score: 66.00
Percent Similarity: 47.27%
Best Local Similarity: 32.73%
Query Match: 21.09%

DB: US-09-727-892A-99 (1-58) x US-09-103-840A-2 (1-4403765)
Length: 4403765
Matches: 18
Conservative: 8
Mismatch: 19
Indels: 10
Gaps: 2

```
OY 7 ThrValleuLeuTYrCysAspGluIle-----LysGlyHisPheProHisGlnIleSer 24
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 55972 ACGGTGCCCTGATGCCGTGAGTTATTGAGCCAAAGCAAGTTCCTACAAATCAAGC 56031
OY 25 MetPheGluAsp-----LeuTYRAspAlaLysValValTyr 36
    |||
    |||||
DB 56032 TGGATCGAAACCGACAGCAGCGACACCGCAGCTGCGCTACGACGACAGATCCGGGTG 56091
OY 37 SerTYrTYrGluTYrAsnLeuPheThrLysLysTYrAlaTyrIle 51
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 56092 CGGTATATGAGATCCGGTGTCTGCTGAGATCTATCACTACCTG 56136
RESULT 2
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 3.56e+04 Length: 4411529
Score: 66.00 Matches: 18
Percent Similarity: 47.27% Conservative: 8
Best Local Similarity: 32.73% Mismatches: 19
Query Match: 21.09% Indels: 10
DB: 4 Gaps: 2

US-09-727-892A-99 (1-58) x US-09-103-840A-1 (1-4411529)
OY 7 ThrValleuLeuTYrCysAspGluIle-----LysGlyHisPheProHisGlnIleSer 24
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 56030 ACGGTGCCCTGATGCCGTGAGTTATTGAGCCAAAGCAAGTTCCTACAAATCAAGC 56089
OY 25 MetPheGluAsp-----LeuTYRAspAlaLysValValTyr 36
    |||
    |||||
DB 56090 TGGATCGAAACCGACAGCAGCGACACCGCAGCTGCGCTACGACGACAGATCCGGGTG 56149
OY 37 SerTYrTYrGluTYrAsnLeuPheThrLysLysTYrAlaTyrIle 51
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 56150 CGGTATATGAGATCCGGTGTCTGCTGAGATCTATCACTACCTG 56194
RESULT 3
US-09-112-096-28
; Sequence 28, Application US/09112096
; Patent No. 6194152
; GENERAL INFORMATION:
; APPLICANT: Reiner Laus
; APPLICANT: Michael H. Shapiro
; APPLICANT: Larisa Tsavaler
; TITLE OF INVENTION: Prostate Tumor Polynucleotide and
; FILE REFERENCE: 7636-0015.30
; CURRENT APPLICATION NUMBER: US/09/112.096
; CURRENT FILING DATE: 1998-07-09
; EARLIER APPLICATION NUMBER: 60/056,110
; EARLIER FILING DATE: 1997-08-20
; NUMBER OF SEQ ID NOS: 29
```

```
SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 3848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-112-096-28

Alignment Scores:
Pred. No.: 19 Length: 3848
Score: 62.00 Matches: 13
Percent Similarity: 52.46% Conservative: 19
Best Local Similarity: 21.31% Mismatches: 15
Query Match: 19.81% Indels: 14
DB: 4 Gaps: 2

US-09-727-892A-99 (1-58) x US-09-112-096-28 (1-3848)
OY 9 LeuLeuTYrCysAspGluIleLysGlyHisPheProHisGlnIleSerMetPheGluAsp 28
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 524 GTCTCTTCTGTGATGAGTGAAGTGAACAGTGTGACTAAATGGGGTGAATTTTACTGAC 583
OY 29 LeuTYRAsp-----AlaLysValValTyrSer 37
    |||||
DB 584 CTGTGAATGTGATGAGACACGCTGGGCTTTTACTTCAATGACGAGATTTGTTCCG 643
OY 38 TyrTYrGluTYrAsn-----LeuPheThrLysLysTYrAlaTyrIleLeuGluTYr 54
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 644 CTCACCTCTGTAAATAAAGCTCTTGTATCTGACGACGATTTTCTGTGACTGAC 703
OY 55 Ile 55
DB 704 ATT 706
RESULT 4
US-09-112-096-14
; Sequence 14, Application US/09112096
; Patent No. 6194152
; GENERAL INFORMATION:
; APPLICANT: Reiner Laus
; APPLICANT: Michael H. Shapiro
; APPLICANT: Larisa Tsavaler
; TITLE OF INVENTION: Prostate Tumor Polynucleotide and
; FILE REFERENCE: 7636-0015.30
; CURRENT APPLICATION NUMBER: US/09/112.096
; CURRENT FILING DATE: 1998-07-09
; EARLIER APPLICATION NUMBER: 60/056,110
; EARLIER FILING DATE: 1997-08-20
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 5668
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-112-096-14

Alignment Scores:
Pred. No.: 30.9 Length: 5668
Score: 62.00 Matches: 13
Percent Similarity: 52.46% Conservative: 19
Best Local Similarity: 21.31% Mismatches: 15
Query Match: 19.81% Indels: 14
DB: 4 Gaps: 2

US-09-727-892A-99 (1-58) x US-09-112-096-14 (1-5668)
OY 9 LeuLeuTYrCysAspGluIleLysGlyHisPheProHisGlnIleSerMetPheGluAsp 28
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2344 GTCTCTTCTGTGATGAGTGAAGTGAACAGTGTGACTAAATGGGGTGAATTTTACTGAC 2403
OY 29 LeuTYRAsp-----AlaLysValValTyrSer 37
    |||||
DB 2404 CTGTGAATGTGATGAGACACGCTGGGCTTTTACTTCAATGACGAGATTTGTTCCG 2463
```

QY 38 TyrTyrGluTyrAsn-----LeuphenhrLysTyrAlaTyrIleIleGluTyr 54
::: |||::: |||::: |||
Db 2464 CTCACGCTGTCTATTAAGCTCTTGTATCTGACGACGATCTTCTGCTGACCTAC 2523
QY 55 Ile 55
|||
Db 2524 ATT 2526
|||
RESULT 5
US-07-885-970A-25
; Sequence 25, Application US/07885970A
; Patent No. 5495070
; GENERAL INFORMATION:
; APPLICANT: John, Maliyakal E.
; TITLE OF INVENTION: GENETICALLY ENGINEERING COTTON
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Nicholas J. Seay, Quarles & Brady
; STREET: P.O. Box 2113, First Wisconsin Plaza
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/885,970A
; FILING DATE: 19920518
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/617,239
; FILING DATE: 21-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/253,243
; FILING DATE: 04-OCT-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27,386
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 283-2478
; TELEFAX: (608) 251-5139
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1984 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Gossypium barbadense
; STRAIN: Sea Island
; IMMEDIATE SOURCE:
; LIBRARY: EMBL ST
; CLONE: SIH6
; US-07-885-970A-25
Alignment Scores:
Pred. No.: 11.5 Length: 1984
Score: 61.00 Matches: 22
Percent Similarity: 50.85% Conservative: 8
Best Local Similarity: 37.29% Mismatches: 23
Query Match: 19.49% Indels: 6
DB: 1 Gaps: 3
US-09-727-892a-99 (1-58) x US-07-885-970A-25 (1-1984)
QY 3 ArgLysTyrLysThrValLeuLeuTyrCys-----AspGluIleLysGlyHisPhe 19

Db 1668 AGACCTACAGTCTTACTTATACATATATGATCTTTCAGACATATATATTTT 1727
||| ||| |||::: ||| ||| |||
QY 20 ProHsgInLiesrMetPheGluAspLeuTyrAspAlaLysValIleTyrSerTyr 39
|||::: |||::: |||::: |||::: |||::: |||
Db 1728 CTTTGTGAGACACTCTTTT-----TTGTAT--TTGAAGAAATTATTTGTTATTTT 1778
|||::: |||::: |||::: |||::: |||::: |||
QY 40 GluTyrAsnLeuphenhrLysTyrAlaTyrIleIleGluTyrIleLysGluIle 58
::: |||::: |||::: |||::: |||
Db 1779 GTTGGAAATATGTTTGTGTGATTTATTCATATATATATATTAAGTAATT 1835
|||::: |||::: |||::: |||
RESULT 6
US-08-298-687A-25
; Sequence 25, Application US/08298687A
; Patent No. 5521078
; GENERAL INFORMATION:
; APPLICANT: John, Maliyakal E.
; TITLE OF INVENTION: GENETICALLY ENGINEERING COTTON
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Nicholas J. Seay, Quarles & Brady
; STREET: P.O. Box 2113, First Wisconsin Plaza
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/298,687A
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/617,239
; FILING DATE: 21-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/253,243
; FILING DATE: 04-OCT-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27,386
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 283-2478
; TELEFAX: (608) 251-5139
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1985 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Gossypium barbadense
; STRAIN: Sea Island
; IMMEDIATE SOURCE:
; LIBRARY: EMBL-ST
; CLONE: SIH6
; US-08-298-687A-25
Alignment Scores:
Pred. No.: 11.5 Length: 1985
Score: 61.00 Matches: 22
Percent Similarity: 50.85% Conservative: 8
Best Local Similarity: 37.29% Mismatches: 23
Query Match: 19.49% Indels: 6
DB: 1 Gaps: 3

```

US-09-727-892A-99 (1-58) x US-08-298-687A-25 (1-1985)
OY      3  ArglysTyrLysThrValLeuLeuTyrCys-----AspGluIleLysGlyHisPhe 19
Db      1669  AGCGCCACGTTACTTTACCATTAATGTATCTTGGAGACATTAAATATATTTT 1728
OY      20  ProHisGluIleSerMetPheGluAspLeuTyrAspAlaLysValValTyrSetTyr 39
Db      1729  CCTTTGGAGACACCTCTTTT-----TTGTAT--TTGAAGCAATTTATTTGTTATTT 1779
OY      40  GluTyrAsnLeuPheThrLysLysTyrAlaTyrIleIleGluTyrIleLysGluIle 58
Db      1780  GTTTGGAATATCGTTGGTTCGATTTATTCGATTCATATATATATATAAAGTAT 1836

RESULT 7
US-08-298-829-25
; Sequence 25, Application US/08298829
; Patent No. 5620882
; GENERAL INFORMATION:
; APPLICANT: John, Mallyakal E.
; TITLE OF INVENTION: GENETICALLY ENGINEERING COTTON
; TITLE OF INVENTION: PLANTS FOR ALTERED FIBER
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nicholas J. Seay, Quarles & Brady
; STREET: P.O. Box 2113, First Wisconsin Plaza
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/298,829
; FILING DATE: 19-OCT-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/885,970
; FILING DATE: 18-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/617,239
; FILING DATE: 21-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/253,243
; FILING DATE: 04-OCT-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27,386
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 283-2478
; TELEFAX: (608) 251-5139
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1985 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Gossypium barbadense
; STRAIN: Sea Island
; IMMEDIATE SOURCE:
; LIBRARY: EMBL-SI
; CLONE: SIH6
; US-08-298-829-25

Alignment Scores:      11.5      Length:      1985
Pred. No.:

```

```

Score: 61.00 Matches: 22
Percent Similarity: 50.85% Conservative: 8
Best Local Similarity: 37.29% Mismatches: 23
Query Match: 19.49% Indels: 6
DB: 1 Gaps: 3

US-09-727-892a-99 (1-58) x US-08-298-829-25 (1-1985)

Cy 3 ArglystYrLsthrValLeuTYrCys-----AspGluIleLySglYHisphe 19
   ||| ||| ||||| ||||| ||| |||
Db 1669 AGAGCCACGTACTTCACTTAATTTGATTCTTTGGAGACAATATATATATTTTTTTA 1728
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Cy 20 ProHISgNIIeSerMetPheGluAspLeuTYrAspAlaLysValValTySerTYrYr 39
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1729 CCTTTTGAGCACCTCTTTT-----TTGAT--TTGAAGAATTTATTTGTTAAATTT 1779
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Cy 40 GluTYrAsnLeuPheThrLYSLYSrYrAlaTYrIleIleGluTYrIleLySgluile 58
   ::|||::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1780 GTTTGGAAATGTTGGTTGGATTATTCGATTCATATATATATAAAGATAAT 1836
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 8
US-08-947-965-1/c
Sequence 1, Application US/08947965A
Patent No. 6004790
GENERAL INFORMATION:
APPLICANT: Dijkstra, Lubbert
APPLICANT: Dijkstra, Bauke
APPLICANT: Andersen, Carsten
APPLICANT: Osten, Claus von der
TITLE OF INVENTION: Cyclomalto-dextrin Glucanotransferase
FILE OF INVENTION: Variants
PCT REFERENCE: 4285.204-US
CURRENT APPLICATION NUMBER: US/08/947,965A
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 0477/95
EARLIER FILING DATE: 1995-04-21
EARLIER APPLICATION NUMBER: 1173/95
EARLIER FILING DATE: 1995-10-17
EARLIER APPLICATION NUMBER: 1881/95
EARLIER FILING DATE: 1995-11-16
EARLIER APPLICATION NUMBER: PCT/DK96/00179
EARLIER FILING DATE: 1996-04-22
NUMBER OF SEQ ID NOS: 78
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 2133
TYPE: DNA
ORGANISM: Thermoaerobacter sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (82)...(2130)
US-08-947-965-1

Alignment Scores:
Pred. No.: 14.8 Length: 2133
Score: 60.50 Matches: 17
Percent Similarity: 60.42% Conservative: 12
Best Local Similarity: 35.42% Mismatches: 13
Query Match: 19.33% Indels: 6
DB: 3 Gaps: 3

US-09-727-892a-99 (1-58) x US-08-947-965-1 (1-2133)

Cy 8 ValLeuTYrCYaspGluIleLySGLYHisPheProHISGlN-----IleSerMet 25
   |||::|||::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 574 GTAATACCTCATGT--CATATATCTCACTTTCCCCAATGAGGTGCTGTCTG 518
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Cy 26 PheGluAspLeuTYrAspAlaLysValValTYrSer-TYrTYrGluTYrAsnLeuPheTh 45
   ::|||::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 517 ATGAGAGAGATGATGAT-----TTGGTGCAAAGTCATATTAATTAATTTGAG 464
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Cy 45 rLysLYsrYrAlaTYrIleIle 52
   ||| |||

```


Db 463 CATGAGCTGTTGCTATGAGATT 442

RESULT 9

US-08-642-274D-25/c

Sequence 25, Application US/08642274D

Patent No. 6200749

GENERAL INFORMATION:

APPLICANT: Shiloh, Yosef

TITLE OF INVENTION: MUTATED FORMS OF THE ATAXIA-TELANGIECTASIA GENE AND METHOD TO

FILE REFERENCE: 229000033

CURRENT APPLICATION NUMBER: US/08/642,274D

CURRENT FILING DATE: 1996-05-03

NUMBER OF SEQ ID NOS: 220

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 25

LENGTH: 476

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: genomic

US-08-642-274D-25

Alignment Scores:

Pred. No.:	3.11	Length:	476
Score:	59.50	Matches:	12
Percent Similarity:	51.11%	Conservative:	11
Best Local Similarity:	26.67%	Mismatches:	21
Query Match:	19.01%	Indels:	1
DB:	4	Gaps:	1

US-09-727-892a-99 (1-58) x US-08-642-274D-25 (1-476)

QY 12 CysAspGluIleLysGlyHisPheProHisGlnIleSerMetPheGluAspLeuTyrAsp 31

Db 407 TGTAAAAAGTTTAAAGAGCATATGTTGTGAGATGCATCCTTTATTACAAAGATATTTCAA 348

QY 32 AAlaYsValValTyrSerTyrTyrGluTyrAsnLeuPheThrLysLysTyrAlaTyrIle 51

Db 347 CAAAAA---AATTATGCTATTATGAAATGATATGACAAAGAGAGACATATCTTA 291

QY 52 IlegluTyrIleLys 56

Db 290 CTAACCTTTTACAAA 276

RESULT 10

US-08-952-014C-25/c

Sequence 25, Application US/08952014C

Patent No. 6265158

GENERAL INFORMATION:

APPLICANT: Shiloh, Yosef

TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE AND ITS

NUMBER OF INVENTION: GENOMIC ORGANIZATION

NUMBER OF SEQUENCES: 91

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kohn & Associates

STREET: 30500 No. 6265158thwestern Hwy., Suite 410

CITY: Farmington Hills

STATE: Michigan

COUNTRY: U.S.

ZIP: 48334

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/952,014C

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Kohn, Kenneth I.

REGISTRATION NUMBER: 30,995

REFERENCE/DOCKET NUMBER: 2290,00028

TELECOMMUNICATION INFORMATION:

TELEPHONE: 810-539-5050

TELEFAX: 810-539-5053

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 476 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-952-014C-25

Alignment Scores:

Pred. No.:	3.11	Length:	476
Score:	59.50	Matches:	12
Percent Similarity:	51.11%	Conservative:	11
Best Local Similarity:	26.67%	Mismatches:	21
Query Match:	19.01%	Indels:	1
DB:	4	Gaps:	1

US-09-727-892a-99 (1-58) x US-08-952-014C-25 (1-476)

QY 12 CysAspGluIleLysGlyHisPheProHisGlnIleSerMetPheGluAspLeuTyrAsp 31

Db 407 TGTAAAAAGTTTAAAGAGCATATGTTGTGAGATGCATCCTTTATTACAAAGATATTTCAA 348

QY 32 AAlaYsValValTyrSerTyrTyrGluTyrAsnLeuPheThrLysLysTyrAlaTyrIle 51

Db 347 CAAAAA---AATTATGCTATTATGAAATGATATGACAAAGAGAGACATATCTTA 291

QY 52 IlegluTyrIleLys 56

Db 290 CTAACCTTTTACAAA 276

RESULT 11

US-08-315-468-3/c

Sequence 3, Application US/08315468

Patent No. 5554534

GENERAL INFORMATION:

APPLICANT: Michaels, Tracy Ellis

APPLICANT: Foncerada, Luis

APPLICANT: Narva, Kenneth E.

TITLE OF INVENTION: Process for Controlling Scarab Pests

TITLE OF INVENTION: with Bacillus thuringiensis Isolates

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: David R. Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/315,468

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/014,941

FILING DATE: 01 FEB 1993

APPLICATION NUMBER: 07/828,430

FILING DATE: 30-JAN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/808,316

FILING DATE: 16-DEC-1991

ATTORNEY/AGENT INFORMATION:

NAME: Saliwanchik, David R.

REGISTRATION NUMBER: 31,794

REFERENCE/DOCKET NUMBER: NA73.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3507 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: kumamotoensis
INDIVIDUAL ISOLATE: 50C
IMMEDIATE SOURCE:
LIBRARY: LambdaEM-11(tm) library of L. Fonceerrada
CLONE: 50C(b)
US-08-315-468-3

Alignment Scores:
Pred. No.: 38.7 Length: 3507
Score: 59.50 Matches: 10
Percent Similarity: 67.65% Conservative: 13
Best Local Similarity: 29.41% Mismatches: 10
Query Match: 19.01% Indels: 1
Gaps: 1

US-09-727-892a-99 (1-58) x US-08-315-468-3 (1-3507)

QY 23 llesermethegluaspleutytr---AspalalysValValtyrSerTytrGluTytr 41
Db 1904 CTCGTATGTTGAAGAAGCATTTACTGAAATGACTCTGGATATTCATATATTGAAA 1845

QY 42 AsnleuphetrllysTytrAlaTytrlleleleTytrlle 55
Db 1844 TCATTATATGTTAAATCATTTTTCATTTTATATAGAAATATGTC 1803

RESULT 12
US-08-743-637B-173/C
Sequence 173, Application US/08743637B
Patent No. 5994066
GENERAL INFORMATION:
APPLICANT: BERGERON, Michel G.
APPLICANT: PICARD, Francois J.
APPLICANT: OUELLETTE, Marc
APPLICANT: ROY, Paul H.
TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA
TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND
TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED
TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS ...
NUMBER OF SEQUENCES: 273
CORRESPONDENCE ADDRESS:
ADDRESSEE: OUARLES & BRADY
STREET: 411 EAST WISCONSIN AVENUE
CITY: MILWAUKEE
STATE: WISCONSIN
COUNTRY: USA
ZIP: 53202-4497
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/743.637B
FILING DATE: 04-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/526,840
FILING DATE: 11-SEP-1995

ATTORNEY/AGENT INFORMATION:
NAME: BAKER, Jean C.
REGISTRATION NUMBER: 35,433
REFERENCE/DOCKET NUMBER: 850586.90012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (414) 277-5000
TELEFAX: (414) 277-5591
INFORMATION FOR SEQ ID NO: 173:
SEQUENCE CHARACTERISTICS:
LENGTH: 630 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-743-637B-173

Alignment Scores:
Pred. No.: 5.24 Length: 630
Score: 59.00 Matches: 21
Percent Similarity: 46.38% Conservative: 11
Best Local Similarity: 30.43% Mismatches: 13
Query Match: 18.85% Indels: 24
Gaps: 5

US-09-727-892a-99 (1-58) x US-08-743-637B-173 (1-630)

QY 5 TyrlsYthrValleuLeuTytrCysAspGluIleTytrGly-----HispheProHsGln 22
Db 537 TATATCAAGCAGCGATGATTTATGATCTGTGATCAAACTTGTTTTCGTTAGCAG 478

QY 23 lleser-----MethegluaspleutytrAspAlaLysVal 34
Db 477 ATTCCCTCAGCTAATCATGATGCGGCTATATCTTT-----TACACACACAGATT 427

QY 35 -----ValtyrSerTytrGluTytrAsnleuphetr 45
Db 426 AGCAGCTACTATTCGACCATCCCGATTCTTACTCTGGCATATATTAACATCTTTCC 367

QY 46 LyslystYtrAlaTytrlleleleTytr 54
Db 366 -----TATCATATCATATTTACTAT 346

RESULT 13
US-08-526-840B-173/C
Sequence 173, Application US/08526840B
Patent No. 6001564
GENERAL INFORMATION:
APPLICANT: BERGERON, Michel G.
APPLICANT: OUELLETTE, Marc
APPLICANT: ROY, Paul H.
TITLE OF INVENTION: SPECIFIC AND UNIVERSAL PROBES AND
TITLE OF INVENTION: AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND IDENTIFY
TITLE OF INVENTION: COMMON BACTERIAL PATHOGENS AND ANTIBIOTIC RESISTANCE GENES
TITLE OF INVENTION: FROM CLINICAL SPECIMENS FOR ROUTINE DIAGNOSIS IN ...
NUMBER OF SEQUENCES: 177
CORRESPONDENCE ADDRESS:
ADDRESSEE: OUARLES & BRADY
STREET: 411 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin
COUNTRY: USA
ZIP: 53202-4497
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/526,840B
FILING DATE: 11-SEP-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/304,732

FILING DATE: 12-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: BAKER, Jean C.
REGISTRATION NUMBER: 35,433
REFERENCE/DOCKET NUMBER: 850586,90012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (414) 277-5000
TELEFAX: (414) 277-5591
INFORMATION FOR SEQ ID NO: 173:
SEQUENCE CHARACTERISTICS:
LENGTH: 630 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-526-840B-173

Alignment Scores:
Pred. No.: 5.24 Length: 630
Score: 59.00 Matches: 21
Percent Similarity: 46.38% Conservative: 11
Best Local Similarity: 30.43% Mismatches: 13
Query Match: 18.85% Indels: 24
DB: 3 Gaps: 5

US-09-727-892A-99 (1-58) x US-08-526-840B-173 (1-630)

Oy 5 TyrLysThrValLeuLeuTyrCysAspGluIleLysGly-----HisPheProHisGln 22
Db 537 TATATCAAGACAGCTGATTTATTTATCTTGATCAATCTTTGTTTATTTTCCTTAGCAGG 478
Oy 23 IleSer-----MetPheGluAspLeuTyrAspAlaLysVal 34
Db 477 ATTTCCTCCAGCTACATGATGATGGCGCTATATCTT-----TACACACAGAAAT 427
Oy 35 -----ValTyrSerTyrTyrGluTyrAsnLeuPheThr 45
Db 426 AGCAGCTACTATTGACACCATCCCGATTTTACTCTCGCATTAATACAAATCTTTCC 367
Oy 46 LysLysTyrAlaTyrIleIleGluTyr 54
Db 366 -----TATCCATCATCATTAACCTAT 346

RESULT 14
US-09-255-893-1
Sequence 1, Application US/09255893A
Patent No. 6008344
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett
TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2 GROUP IV EXPRESSION
FILE REFERENCE: RRS-0055
CURRENT APPLICATION NUMBER: US/09/255,893A
CURRENT FILING DATE: 1999-02-23
NUMBER OF SEQ ID NOS: 47
SEQ ID NO 1
LENGTH: 2875
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (139)..(2388)
US-09-255-893-1

Alignment Scores:
Pred. No.: 35.6 Length: 2875
Score: 59.00 Matches: 14
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 58.33% Mismatches: 6
Query Match: 18.85% Indels: 2
DB: 3 Gaps: 1

US-09-727-892A-99 (1-58) x US-09-255-893-1 (1-2875)

Oy 35 ValTyrSerTyrTyrGluTyrAsnLeuPheThrLysTyrAlaTyrIleIleGluTyr 54
Db 2774 ATTATATCATACATGACAGACACACTATTTTATTTATATGATATATA-----TAC 2827
Oy 55 IleLysGluIle 58
Db 2828 ATACATGAATA 2839

RESULT 15
US-08-930-996A-1
Sequence 1, Application US/08930996A
Patent No. 6100449
GENERAL INFORMATION:
APPLICANT: Fluhr, Robert
APPLICANT: Eshed, Yuval
APPLICANT: ORI, Naomi
APPLICANT: PARAN, Ilan
APPLICANT: ZAMIR, Daniel
TITLE OF INVENTION: A GENE FAMILY FROM THE I2 FUSARIUM RESISTANCE
LOCUS OF TOMATO AND USE THEREOF FOR TRANSPORTATION AND
SELECTIVE BREEDING OF TOMATO AND RELATED PLANTS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEWMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/930,996A
FILING DATE: 09-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/05272
FILING DATE: 15-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 113,373
FILING DATE: 13-APR-1995
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4946 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 299..3958
US-08-930-996A-1

Alignment Scores:
Pred. No.: 70.6 Length: 4946
Score: 59.00 Matches: 12
Percent Similarity: 54.29% Conservative: 7
Best Local Similarity: 34.29% Mismatches: 16
Query Match: 18.85% Indels: 0
DB: 3 Gaps: 0

US-09-727-892A-99 (1-58) x US-08-930-996A-1 (1-4946)

Oy 9 LeuLeuTyrCysAspGluIleLysGlyHisPheProHisGlnIleSerMetPheGluAsp 28
Db 3374 CTGACTTATTTGCGAAGATAGACAGAGAAATGCGCTTCATTTTCAATATCTCATATC 3433
Oy 29 LeuTyrAspAlaLysValValTyrSerTyrTyrGluTyrAsnLeu 43
Db 3434 AGATATTGCAAGAAACCTGTGAATGCGCGAAGAGAGTGCGATTTA 3478

Wed Nov 6 12:59:58 2002

us-09-727-892a-99c Iran. rni

Page 8

```
Search completed: November 5, 2002, 05:01:18
Job time : 6297 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 5, 2002, 03:12:08 ; Search time 1728 Seconds

(without alignments)
453.023 Million cell updates/sec

Title: US-09-727-892a-99

Perfect score: 313

Sequence: 1 MERRKYVLYTCDEIKGHF.....YEVNLFTRKAYITIEYKEI 58

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 13736207 segs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+2n.model -DEV=xlh
-O/cgn2.1/USPTO/US09727892/transcript_01112002.185643.3873/app.query.fasta_1.199
-DB=EST -QFMT=fastp -SUFFIX=transcript -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human4.0.cdi -LIST=45
-LOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MTLEN=0 -MAXLEN=200000000
-USER=US09727892.ecgn.1.1.763.transcript_01112002.185643.3873 -NCTP=6 -ICPU=3
-NO_XUPXT -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_estbhum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estlro:*
8: em_hic:*
9: gb_estl1:*
10: gb_estl2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82.5	26.4	1027	12	CNS06H35
2	76	24.3	831	12	BH382856

3	75	24.0	565	9	AA656172
4	74.5	23.8	922	10	BG028123
5	71.5	22.8	452	10	B1039914
6	70	22.4	375	9	AA255939
7	70	22.4	454	9	AA391244
8	70	22.4	458	10	R92556
9	70	22.4	593	10	H11366
10	70	22.4	621	10	W06896
11	70	22.4	695	9	BS527932
12	69.5	22.2	523	10	B1041289
13	69.5	22.2	608	9	AW393177
14	69	22.0	590	9	AV860055
15	69	22.0	673	10	B134055
16	69	22.0	1428	12	B13405
17	68.5	21.9	692	12	BH121163
18	68.5	21.9	839	12	BH601582
19	68.5	21.9	973	9	A1068370
20	68	21.7	343	9	A1872988
21	68	21.7	618	10	B1505912
22	68	21.7	810	12	BH488035
23	67.5	21.6	538	10	B1921487
24	67.5	21.6	739	12	A0257277
25	67	21.4	396	12	A0093101
26	67	21.4	436	9	AA676845
27	67	21.4	437	9	AA043694
28	67	21.4	570	9	AV860333
29	66.5	21.2	565	9	A1848036
30	66.5	21.2	569	9	AW393171
31	66	21.1	478	9	AW106487
32	66	21.1	501	10	BM275504
33	66	21.1	556	9	AA553855
34	66	21.1	579	10	BM275862
35	66	21.1	644	12	AO899545
36	66	21.1	720	10	BM164432
37	66	21.1	880	12	AA693556
38	65.5	20.9	432	12	BH047620
39	65.5	20.9	376	10	BM163188
40	65.5	20.9	584	9	AW393172
41	65.5	20.9	746	10	BG622859
42	65	20.8	240	12	A928839
43	65	20.8	288	9	BB394906
44	65	20.8	297	9	A1167082
45	65	20.8	304	9	AA322724

ALIGNMENTS

RESULT 1	CNS06H35	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	TITLE	JOURNAL	MEDLINE	REFERENCE	AUTHORS	
1	CNS06H35	1027 bp	DNA	linear	GSS	17-JUN-2001	T3 end of clone AS0A009E02 of library AS0A from strain CLIB 533	AL398487	GI:12152025	GSS.	Saccharomyces bayanus.	Saccharomyces bayanus.	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.	1 (bases 1 to 1027)	Bon,E., Neuveglise,C., Casaregola,S., Artiguenave,F., Wincker,P., Aigle,M. and Durans,P.
2	20584715	Genomic exploration of the hemiascomycetous yeasts: 5.	Saccharomyces bayanus var. uvarum	FEBS Lett. 487 (1), 37-41 (2000)	2 (bases 1 to 1027)	Soucier,J.L., Aigle,M., Artiguenave,F., Blandin,G., Boitlin-Pukharra,M., Bon,E., Brotier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durans,P., Lepingle,A., Liorette,B., Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S., Saurin,M., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,P. and Weissenbach,J.									

TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies

JOURNAL FEBS Lett. 487 (1), 3-12 (2000)

MEDLINE 20584711

REFERENCE 3 (bases 1 to 1027)

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremlieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail: seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT This GSS is part of a random genomic sequencing program of thirteen yeast species: *Saccharomyces bayanus* var. *uvorum*, *Saccharomyces exiguus*, *Saccharomyces servazzii*, *Zygosaccharomyces rouxii*, *Saccharomyces kluyveri*, *Kluyveromyces thermotolerans*, *Kluyveromyces lactis* var. *lactis*, *Kluyveromyces marxianus* var. *marxianus*, *Pichia angusta*, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*, *Candida tropicalis* and *Yarrowia lipolytica*. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.

FEATURES

source

1..1027

/organism="Saccharomyces bayanus"

/strain="ClIB 533"

/variety="uvorum"

/db_xref="taxon:4931"

/clone="AS0A009E02"

/clone_1ib="AS0AA"

/note="end : T3"

<296..>661

/note="similar to *Saccharomyces cerevisiae* ORF YDR357c [hypothetical protein]"

/evidence="not_experimental"

complement(<901..>1014)

/note="similar to *Saccharomyces cerevisiae* ORF YDR356w [NUP1; spindle pole body component]"

/evidence="not_experimental"

BASE COUNT 336 a 181 c 190 g 318 t 2 others

ORIGIN

Alignment Scores:

Pred. No.:	1.01	Length:	1027
Score:	82.50	Matches:	17
Percent Similarity:	47.83%	Conservative:	16
Best Local Similarity:	24.64%	Mismatches:	19
Query Match:	26.36%	Indels:	17
DB:	12	Gaps:	2

US-09-727-892a-99 (1-58) x CNS06H35 (1-1027)

QY 4 LysTYLysThValleuLeuTYrCYsAspGluIleLysGlyHisPhePro-----His 21

Db 509 AAATTCGAGGAGCTTCACAGAAATCGCATGATAGAACACATTTTGAGATGTTAAAC 568

QY 22 GlnIleSerMetPheGluAspLeuTYrAspAlaLysValTYrSerTYrGluTYr 41

Db 569 CAATAGCAATGATACGATCTTCAAGAGAGAGATAGCAGTACCGTCGACGATTC 628

QY 42 AsnLeuPheThrLys----- 46

Db 629 AACGATTAAAGAGGCTGTTAATAAATAAGTAGTACCTATATACAAATACGCAAC 688

QY 47 LysTYrAlaTYrIleGluTYrIle 55

Db 689 GTATATACACATATATGCGAATATGTG 715

RESULT 2

LOCUS BH382856/c 831 bp DNA linear GSS 10-DEC-2001

DEFINITION AG-ND-17713.TR ND-TAM Anopheles gambiae genomic clone AG-ND-17713,

ACCESSION BH382856

VERSION BH382856.1 GI:17328998

KEYWORDS GSS.

SOURCE African malaria mosquito.

ORGANISM Anopheles gambiae

TITLE Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.

REFERENCE 1 (bases 1 to 831)

AUTHORS Shetty,J., Malek,J., Koo,H., Collins,F., Gardner,M. and Loftus,B.J.

TITLE Direct Submission of BAC-end sequences from *Anopheles gambiae*

JOURNAL Unpublished (2001)

COMMENT Other-GSSs: AG-ND-17713.TF

Contact: Brendan J Loftus

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208

Fax: 301 838 3543

Email: bjoftus@tifg.org

This clone is from an A. gambiae BAC library (ND-TAM) provided by F.H. Collins and sequenced by The Institute for Genomic Research (TIGR). The BAC library was generated from A. gambiae PEST strain DNA. All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest.

Seq primer: M13 Rev

Class: BAC ends.

FEATURES

source

1..831

/organism="Anopheles gambiae"

/strain="PEST"

/db_xref="taxon:7165"

/clone="AG-ND-17713"

/clone_1ib="ND-TAM"

/note="Vector: pECBAC1, Site_1: HindIII"

BASE COUNT 292 a 146 c 138 g 255 t

ORIGIN

Alignment Scores:

Pred. No.:	5.47	Length:	831
Score:	76.00	Matches:	16
Percent Similarity:	59.57%	Conservative:	12
Best Local Similarity:	34.04%	Mismatches:	15
Query Match:	24.28%	Indels:	4
DB:	12	Gaps:	1

US-09-727-892a-99 (1-58) x BH382856 (1-831)

QY 12 CysAspGluIleLysGlyHisPheProHisGlnIleSerMetPheGluAspLeuTYrAsp 31

Db 661 TGTGGGAAATGGAAATATTTTCACCTGCTTAAACAGCAGCAATTTGACGAC 602

QY 32 AlaLysValTYrSerTYrTYrGluTYrAsnLeuPheThrLysLys----- 47

Db 601 AAGAGCTTACATGAGCTGATATAAACAATGATTTACAAAAAATAGTGCATCTT 542

QY 48 TYrAlaTYrIleGluTYr 54

Db 541 AGATCTTGATCATACGCTAC 521

RESULT 3

LOCUS AA656172 565 bp mRNA linear EST 04-NOV-1997

DEFINITION vs50fil.r1 StrataGene mouse Tcell 937311 Mus musculus cDNA clone IMAGE:1149741 5' similar to gb:S66915_cds1 Arp SYNTHASE GAMMA CHAIN, MITOCHONDRIAL, PRECURSOR (HUMAN);, mRNA sequence.

ACCESSION AA656172

VERSION AA656172.1 GI:2592326

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=MR4&t2=MR4-NT0141-310101-206-d06&t3=2001-01-31&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 452.
Location/Qualifiers

FEATURES

Source

1..452
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="NT0141"
/dev_stage="Adult"

/note="Organ: nervous tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196/716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 109 a 125 c 91 g 127 t
ORIGIN

Alignment Scores:

Pred. No.: 10 Length: 452
Score: 71.50 Matches: 16
Percent Similarity: 57.78% Conservative: 16
Best Local Similarity: 35.56% Mismatches: 16
Query Match: 22.84% Indels: 3
Gaps: 1

US-09-727-892A-99 (1-58) x BI039914 (1-452)

QY 16 LysGlyHisPheProHISGlnIleSerMetPheGluAspLeuTyrAspAlaLysValAla 35

Db 316 AAGGGGCAAGTTTGGGCGAGGTTCTCGAGTTCGAAATTCGCCGCGAAGGCTTGA 257

QY 36 TyrSerTyr-----TyrGluTyrAsnLeuPheThrLysTyrAlaTyrIlelle 52

Db 256 TATGCTACACAGCAATTCACAAATGATGTGATGACACAGAAAGAGCTTATGTGCAC 197

QY 53 GluTyrIleLysGlu 57

Db 196 AAGAGGTATATGGA 182

RESULT 6 AA255939 375 bp mRNA linear EST 13-AUG-1997

LOCUS zs29a01.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:686568 3'

DEFINITION similar to contains Alu repetitive element; mRNA sequence.

ACCESSION AA255939

VERSION AA255939.1 GI:1891578

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 375)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
NATIONAL Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-f@mail.nih.gov

This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 844 Std Error: 0.00
Seq primer: -41ml3 fwd. ET from Amersham
High quality sequence stop: 374.
Location/Qualifiers

FEATURES

Source

1..375
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="IMAGE:686568"
/clone_1lb="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"

/db_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI). Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - Oligo(dT) primer
15'-TGTTCACATCTGAGAGTGAGGAGCGCCCTCATTTTCTTTTCTTTT-3'
1. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 112 a 85 c 80 g 98 t
ORIGIN

Alignment Scores:

Pred. No.: 12.5 Length: 375
Score: 70.00 Matches: 20
Percent Similarity: 50.94% Conservative: 7
Best Local Similarity: 37.74% Mismatches: 24
Query Match: 22.36% Indels: 2
Gaps: 0

US-09-727-892A-99 (1-58) x AA255939 (1-375)

QY 7 ThrValLeuLeuTyrCysAspGluIleLysGlyHisPheProHISGlnIleSerMetPhe 26

Db 130 ACACTACTCCCTGACTGTGATCTTCAGAAATTTGTTACATCATACATCAAGGCTC 189

QY 27 GluAspLeuTyrAspAlaLysValTyrSerTyrTyr--GluTyrAsnLeuPheThrL 46

Db 190 TCGTCAGTCATGATGATGCTCCATGCTACTCATCATCATCATCATCATCATCAT 249

QY 46 yLysTyrAlaTyrIlelleGluTyrIleLysGlu 57

Db 250 TTAACATACATGCTGTGGAAGGCTGATGAGAGA 284

RESULT 7 AW391244 454 bp mRNA linear EST 04-FEB-2000

LOCUS QY0-ST0213-021299-062-d03 ST0213 Homo sapiens cDNA, mRNA sequence.

ACCESSION AW391244

VERSION AW391244.1 GI:6896007

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 454)
HCGP http://www.ludwig.org.br/ORESTES.
The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
JOURNAL Contact: Simpson A.J.G.
COMMENT Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 1242
 High quality sequence stops: 495
 Source: IMAGE Consortium, LNL
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 1242 Std Error: 0.00
 Seq primer: M13Rev
 High quality sequence stop: 495.

FEATURES

```

/organism="Homo sapiens"
/db_xref="GDB:420289"
/db_xref="taxon:9606"
/clone="IMAGE:47748"
/clone_1lb="Soares Infant brain INIB"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lambda I; Site:1: Not I; Site:2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5], AACGCGAAGATTGCGGCCCGCCAGCAATTTTTTTTTTTT 3'; double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lambda I vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Patima Bonaldo."

```

Alignment Scores:

Pred. No.:	21.9	length:	593
Score:	70.00	Matches:	13
Percent Similarity:	53.85%	Conservative:	15
Best Local Similarity:	25.00%	Mismatches:	20
Query Match:	22.36%	Indels:	4
DB:	10	Gaps:	1

US-09-727-892A-99 (1-58) x H11366 (1-593)

7 ThrValLeuLeuTyrCysaspGluIleLysGlyHisPheProHisGlnIleSerMetPhe 26

Db 256 A C A C T A C T G T T C A T C T T A C C T C T T T G T C T G G T C C A C T C C C C C A T C A T A C T C A G T C T T C 19

2 / GUASPLeuIYrAspAlaLysValValTYrSerTYrTYrGluTYrAsnLeuPheThrLys 46

[illegible][illegible]

RECITE 10

W06896/c
LOCUS
W06896
621 bp
mpna
140000
EST 25-7ABD-1006

DEFINITION
zab89g10.r1 soares_tetal_lung_NbHL19W Homo sapiens cDNA clone
IMAGE:299778 5' mRNA sequence.

ACCESSION	W06896
VERSION	W06896.1 GI:1280917

SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, F., Waterston

TITLE	EST	PROJECT
The Washu-Merck		

Unpublished (1995).
Contact: Wilson R.

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: mob.REGA+ET
High quality sequence stop: 266.

FEATURES

```

/organism="Homo sapiens"
/db_xref="GDB:1244702"
/db_xref="taxon:9606"
/clone="IMAGE:299778"
/clone_11b="Soares_fetal_lung_NbHL19"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/motif="Organ: lung; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [3'-GTATTCACATCTGAAGTGGAGCGGCCGCAATTTTATTTTATTTT-3'] double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Patima Bonaldo. This library was constructed from the same fetus as the fetal heart library, Soares fetal heart NbHL19."

```

ORIGIN

Alignment Scores:	
Prod. No.:	23.2
Score:	70.00
Percent Similarity:	53.85%
Best Local Similarity:	25.00%
Query Match:	22.36%
Length:	621
Matches:	13
Conservative:	15
Mismatches:	20
Indels:	4
Gaps:	1

US-09-727-892A-99 (1-58) x W06896 (1-621,

7 ThrValLeuLeuTyrCysaspGluIleLysGlyHisPheProHisGlnIleSerMetPhe 26

293 ACACCTACTGTTTCATCTCTTACCTCTTTTGTCTCTCCACTCCCTCCATCATTACCTCAGTCTTC 234

```

21 61488pdeu1y1nspralabysva1val1y1se1y11y1g1u1y1nsu1deu1ne1m1u1ys 40
    :: ::||| :: :: |||::: |||||

```

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

SECRET 11

BB527932/c
LOCUS
BB527932
695 bp
mRNA
linear
EST 25-OCT-2001

musculus cDNA clone D930043P22 3', mRNA sequence

REVISION	BB527932
VERSION	BB527932.2
	GI:16444564

SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus

REFERENCE	1	(bases 1 to 695)
AUTHORS	Arakawa, T.,	Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.

HIRAMOTO, K., HOLI, F., ISHII, I., ITO, M., KAWAI, J., KOHNO, H.,
M., KOYA, S., MATSUYAMA, T., MATUZAKI, A., NOMURA, K., OHNO, M.,

was cloned into the *NotI* and *BamHI* sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda Φ 101.

```

locus_tag     locus_tag:
/note="Organ: nervous_tumor; Vector: puc18; Site_1: SmaI;
site 2: SmaI. A mini-library was made by cloning products

```


Oy	22	gnliesermtpegluaspleutyaspalsalvalyrisertyr	39
		::: ::: :::	
Db	276	gaamtgcttgccttgcgcacttacttncacagatctggcaatctatgctac	217
Oy	40	glutyrasn-----leuphetrlylsylstyralatyr	50
		::: :::	
Db	216	aattatataatcaatcaagaactcttgactctgaagcactatgggat	172

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 5, 2002, 03:12:53 ; Search time 217 Seconds

(without alignments)
458.899 Million cell updates/sec

Title: US-09-727-892A-99

Perfect score: 58

Sequence: 1 MERKTVLLCYDEIKGHF.....YEVNLFKKYATIIETIKEL 58

Scoring table:

OLIGO	
Xgapop 60.0 , Ygapext 60.0	
Xgapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1736436 segs, 858457221 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3471287

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODL=frame+P2n.model -DEV=xlh
-Q=C9a2_1/OSPro.spool/US09727892/runat_01112002_185927_4870/app.query.fasta.1.199
-DB=N_Geneseq.032802 -QFWT=fastap -SUFFIX=Naolig.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=45 -MODE=LOCAL
-OUTFMT=ptio -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09727892_ECGN_1.1.0.etruncat_01112002_185927_4870 -NCP0=6 -ICPU=3
-NO_XLPXY -NO_MMAB -LARGEOUTERY -NES_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N_Geneseq_032802:*

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	100.0	177	21 AAA69042	Bacteriophage 44AH
2	58	100.0	1668	21 AAA69013	Bacteriophage 44AH
3	52	89.7	165	21 AAA69045	Bacteriophage 44AH
4	5	12.1	235	21 AAC36623	Eucalyptus grandis
5	7	12.1	378	22 AAI85901	Human polynucleoti
6	7	12.1	509	22 AAS60503	Human cancer agent
7	7	12.1	788	21 AAA01729	Human colon cancer
8	7	12.1	970	20 AAX97705	Extended human sec
9	7	12.1	1200	21 AAC45926	Arabidopsis thalia
10	7	12.1	1754	21 AAC42322	Arabidopsis thalia
11	7	12.1	2163	21 AAF21038	Human adenosine re
12	7	12.1	2163	21 AAA34916	Human adenosine re
13	7	12.1	2200	23 ABL19908	Drosophila melanog
14	7	12.1	2242	23 ABL09205	Drosophila melanog
15	7	12.1	2418	22 AAF64195	Human secreted pro
16	7	12.1	2418	22 AAF32736	Human secreted pro
17	7	12.1	2433	22 AAD09336	Rat PRPDE90 cDNA e
18	7	12.1	2497	11 AAF21044	HindIII S fragment
19	7	12.1	2511	21 AAF21040	Human low adenosin
20	7	12.1	2511	21 AAA34918	Human adenosine re
21	7	12.1	2793	21 AAA70146	Plasmodium falcipa
22	7	12.1	3022	22 AAD09338	Rat PRPDE89 cDNA e
23	7	12.1	3133	22 AAD09340	Rat PRPDE74 cDNA e
24	7	12.1	3186	12 AA014631	Plasmod pTM72 inse
25	7	12.1	3186	17 AAT34375	Plasmod pTM72 (ATC
26	7	12.1	3186	20 AAZ32242	Human glioblastoma
27	7	12.1	3192	23 ABL14027	Drosophila melanog
28	7	12.1	3233	23 ABA88177	pTM72 human gliobl
29	7	12.1	5080	23 ABL09204	Drosophila melanog
30	7	12.1	5536	23 ABL14026	Drosophila melanog
31	7	12.1	6195	24 ABL32591	Human immune syste
32	7	12.1	6219	24 ABL32866	Human immune syste
33	7	12.1	6447	24 AAS63324	Chemically pretrea
34	7	12.1	6447	23 ABL15218	Drosophila melanog
35	7	12.1	6511	23 ABL22283	Drosophila melanog
36	7	12.1	6519	22 AAK85381	Human immune/haema
37	7	12.1	6581	22 AAK85383	Human immune/haema
38	7	12.1	7275	23 ABL22879	Drosophila melanog
39	7	12.1	9001	22 AAK84675	Human immune/haema
40	7	12.1	11424	23 ABL22878	Drosophila melanog
41	7	12.1	11641	22 ABA14848	Human nervous syst
42	7	12.1	11641	22 AAK85387	Human immune/haema
43	7	12.1	12904	22 AAK85382	Human immune/haema
44	7	12.1	13293	23 ABL22282	Drosophila melanog
45	7	12.1	13814	24 ABL33192	Human immune syste

ALIGNMENTS

RESULT 1
AAA69042
ID AAA69042 standard; DNA; 177 BP.

AAA69042:
27-0CT-2000 (first entry)

Bacteriophage 44AHUD nucleotide sequence 44AHUDORF025.

Bacteriophage; antimicrobial; genome; identification; antibacterial;
bacterial growth inhibition; bacterial infection; ds.

Bacteriophage 44AHUD.

WO200032825-A2.

08-JUN-2000.

```
XX 03-DEC-1999; 99WO-IB02040.
PF
XX 03-DEC-1998; 98US-0110992.
PR 03-JUN-1999; 99US-0326144.
PR 28-SEP-1999; 99US-0407804.
PR 30-SEP-1999; 99US-0157218.
PR 01-DEC-1999; 99US-0168777.
PR 02-DEC-1999; 99US-0454252.
XX
XX (PHAG-) PHAGETECH INC.
PI Pelletier J, Gros P, Dubow M;
XX WPI; 2000-412361/35.
DR P-PSDB; AAB16557.
XX
XX Identifying a bacteriophage coding region for treating bacterial
PT infections comprises identifying a nucleic acid encoding a product that
PT inhibits bacteria when a bacteriophage infects a bacterium
XX
XX Example 9; Page 278; 456pp; English.
XX
XX The present invention describes a method for identifying a bacteriophage
CC coding region encoding a product active on an essential bacterial
CC target. The method comprises identifying a nucleic acid sequence encoding
CC a gene product that provides a bacteria-inhibiting function when an
CC uncharacterised bacteriophage infects a pathogenic bacterium. The
CC compound active on a target of a bacteriophage inhibitor protein in a
CC bacteria is used to treat or prevent a bacterial infection in an animal.
CC AAB68243 to AAB69442 and AAB16523 to AAB16954 represent bacteriophage
CC nucleotide and protein sequences which are used in the exemplification of
CC the present invention.
XX
XX SQ Sequence 177 BP; 72 A; 22 C; 26 G; 57 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 3.35e-55 Length: 177
XX Score: 58.00 Matches: 58
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 21 Gaps: 0
XX
XX US-09-727-892a-99 (1-58) x AAA69042 (1-177)
XX
XX QY 1 MetGLuArGlySTyTrhValLeuLeuTyrcysAspGluileLysGlyHisPhePro 20
XX |||||||
XX Db 1 ATGGAAAGCTAAATACAAAGCGATTATTATATTGCGATGAGATTAAAGACATTTTCCA 60
XX
XX QY 21 HisGlnIleSerMetPheGluAspLeuTyrcysAspAlaLysValValTyrcSerTyrcGlu 40
XX |||||||
XX Db 61 CATCAATATCTCAATGTTGAAGATTATATGACGCTAAAGTGTATATTCATATTATGAA 120
XX
XX QY 41 TyrAsnLeuPheThrLysLysTyrcAlaTyrcIleIleGluTyrcIleLysGluile 58
XX |||||||
XX Db 121 TATTAACCTGTTCACTAAAAAATACGCTATATCATAGAAATACATTAAAGAGATA 174
XX
XX RESULT 2
XX ID AAA69013 standard; DNA; 1668 BP.
XX
XX AC AAA69013;
XX
XX DT 27-OCT-2000 (first entry)
XX
XX DE Bacteriophage 44AHUD complete genome sequence.
XX
XX KW Bacteriophage; antimicrobial; genome; identification; antibacterial;
XX bacterial growth inhibition; bacterial infection; ds.
XX
XX OS Bacteriophage 44AHUD.
XX
```

```
EN WO200032825-A2.
XX
XX 08-JUN-2000.
XX
XX 03-DEC-1999; 99WO-IB02040.
PF
XX 03-DEC-1998; 98US-0110992.
PR 03-JUN-1999; 99US-0326144.
PR 28-SEP-1999; 99US-0407804.
PR 30-SEP-1999; 99US-0157218.
PR 01-DEC-1999; 99US-0168777.
PR 02-DEC-1999; 99US-0454252.
XX
XX (PHAG-) PHAGETECH INC.
PI Pelletier J, Gros P, Dubow M;
XX WPI; 2000-412361/35.
DR
XX
XX Identifying a bacteriophage coding region for treating bacterial
PT infections comprises identifying a nucleic acid encoding a product that
PT inhibits bacteria when a bacteriophage infects a bacterium
XX
XX Example 9; Page 266-269; 456pp; English.
XX
XX The present invention describes a method for identifying a bacteriophage
CC coding region encoding a product active on an essential bacterial
CC target. The method comprises identifying a nucleic acid sequence encoding
CC a gene product that provides a bacteria-inhibiting function when an
CC uncharacterised bacteriophage infects a pathogenic bacterium. The
CC compound active on a target of a bacteriophage inhibitor protein in a
CC bacteria is used to treat or prevent a bacterial infection in an animal.
CC AAB68243 to AAB69442 and AAB16523 to AAB16954 represent bacteriophage
CC nucleotide and protein sequences which are used in the exemplification of
CC the present invention.
XX
XX SQ Sequence 16668 BP; 6095 A; 2338 C; 2608 G; 5627 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 2.03e-53 Length: 16668
XX Score: 58.00 Matches: 58
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 21 Gaps: 0
XX
XX US-09-727-892a-99 (1-58) x AAA69013 (1-16668)
XX
XX QY 1 MetGLuArGlySTyTrhValLeuLeuTyrcysAspGluileLysGlyHisPhePro 20
XX |||||||
XX Db 15175 ATGGAAAGCTAAATACAAAGCGATTATTATATTGCGATGAGATTAAAGACATTTTCCA 15116
XX
XX QY 21 HisGlnIleSerMetPheGluAspLeuTyrcysAspAlaLysValValTyrcSerTyrcGlu 40
XX |||||||
XX Db 15115 CATCAATATCTCAATGTTGAAGATTATATGACGCTAAAGTGTATATTCATATTATGAA 15056
XX
XX QY 41 TyrAsnLeuPheThrLysLysTyrcAlaTyrcIleIleGluTyrcIleLysGluile 58
XX |||||||
XX Db 15055 TATTAACCTGTTCACTAAAAAATACGCTATATCATAGAAATACATTAAAGAGATA 15002
XX
XX RESULT 3
XX ID AAA69045 standard; DNA; 165 BP.
XX
XX AC AAA69045;
XX
XX DT 27-OCT-2000 (first entry)
XX
XX DE Bacteriophage 44AHUD nucleotide sequence 44HJDRF029.
XX
XX KW Bacteriophage; antimicrobial; genome; identification; antibacterial;
XX bacterial growth inhibition; bacterial infection; ds.
XX
```



```
OS Bacteriophage 44AHJD.
XX
PN WO200032825-A2.
XX
PD 08-JUN-2000.
XX
PF 03-DEC-1999; 99MO-IB02040.
XX
PR 03-DEC-1998; 98US-0110992.
PR 03-JUN-1999; 99US-0326144.
PR 28-SEP-1999; 99US-0407804.
PR 30-SEP-1999; 99US-0157218.
PR 01-DEC-1999; 99US-0168777.
PR 02-DEC-1999; 99US-0454252.
XX
PA (PHAG-) PHAGETECH INC.
XX
PI Pelletier J, Gros P, Dubow M;
XX
DR WPI: 2000-412361/35.
DR P-PSDB; AAB16560.
XX
XX
PT Identifying a bacteriophage coding region for treating bacterial
PT infections comprises identifying a nucleic acid encoding a product that
PT inhibits bacteria when a bacteriophage infects a bacterium
XX
PS Example 9; Page 278; 456pp; English.
XX
XX
CC The present invention describes a method for identifying a bacteriophage
CC coding region encoding a product active on an essential bacterial
CC target. The method comprises identifying a nucleic acid sequence encoding
CC a gene product that provides a bacteria-inhibiting function when an
CC uncharacterised bacteriophage infects a pathogenic bacterium. The
CC compound active on a target of a bacteriophage inhibitor protein in a
CC bacteria is used to treat or prevent a bacterial infection in an animal.
CC AAB68243 to AAB69442 and AAB16523 to AAB16954 represent bacteriophage
CC nucleotide and protein sequences which are used in the exemplification of
CC the present invention.
XX
SQ Sequence 165 BP; 63 A; 21 C; 25 G; 56 T; 0 other;
XX
Alignment Scores:
Pred. No.: 1.19e-48 Length: 165
Score: 52.00 Matches: 52
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.66% Indels: 0
DB: 21 Gaps: 0
XX
US-09-727-892a-99 (1-58) x AAB69045 (1-165)
QY 1 MetLUARGLYSTRYLSTHVALLEULEUTYCYASPSGLIUELYSGLYHISPHEDRO 20
DB 9 ATGGAACGTAAATCAAAACGCTATTATTCATGCGATGAGATTAAAGCATTTCACA 68
QY 21 HiscGlnIleSerMetPheGluAspLeuTYrAspAlaLysValLYrSerTYrTYGlu 40
DB 69 CATCAAAATCTCAATGTTGAAGATTATATGACGCTAAAGTTGATATTCATTATATGAA 128
QY 41 TYrAsnLeuPheThrLYrLYrLYrLYrLYrLYrLYrLYrLYrLYrLYrLYrLYrLYr 52
DB 129 TATAAAGCTGTCACCTAAATAAATACGCGTATATCATATA 164
XX
RESULT 4
ID AAC56623 standard; DNA; 235 BP.
XX
AC AAC56623;
XX
DT 25-JAN-2001 (first entry)
XX
DE Eucalyptus grandis transcription factor DNA sequence #494.
XX
```

```
KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
KW homeodomain zipper; LIM domain; AP2; EREBS; zinc finger domain;
KW type 2 Cys2His2; CCAAT box element; MYB; SS.
XX
OS Eucalyptus grandis.
XX
PN WO200053724-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000MO-US06112.
XX
PR 11-MAR-1999; 99US-0266513.
PR 18-AUG-1999; 99US-0149485.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Wood M, McGrath A, Shenk MA, Glenn M;
XX
DR WPI: 2000-579369/54.
XX
XX
PT New isolated polynucleotide encoding a plant transcription factor for
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
PT having modified gene expression or modified activity of a polypeptide
PT
XX
PS Claim 1; Page 478; 747pp; English.
XX
XX
CC The present invention relates to novel plant transcription factors from
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding
CC sequence for one such transcription factor. The transcription factor may
CC be used to produce a plant having modified gene expression such as a
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or
CC mahogany species or to modify the activity of a polypeptide in a plant.
CC The transcription factors of the present invention are members from the
CC following families of regulatory proteins: bZIP, bZIP family of G-box
CC binding factors; basic helix-loop-helix zipper, LIM domain, AP2
CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain,
CC and EREBS, zinc finger domains of type 2 Cys2His2, CCAAT box elements
CC and MYB.
XX
SQ Sequence 235 BP; 32 A; 118 C; 52 G; 33 T; 0 other;
XX
Alignment Scores:
Pred. No.: 34.2 Length: 235
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 21 Gaps: 0
XX
US-09-727-892a-99 (1-58) x AAC56623 (1-235)
QY 37 SerTYrTYrGluTYrAsnLeu 43
DB 64 TCCTACTACGAATACAACTC 84
XX
RESULT 5
ID AA185901 standard; cDNA; 373 BP.
XX
AC AA185901;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 5961.
XX
DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW
```

KW nervous system disorders; arthritis; inflammation; ss.
XX Homo sapiens.
OS
XX WO200164835-A2.
PN
XX
XX 07-SEP-2001.
PD
XX
XX 26-FEB-2001; 2001WO-US04927.
PF
XX
XX 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
PX
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Drmanac RT;
PI
XX
XX WPI: 2001-514838/56.
DR
XX
XX P-PSDB; AAO05970.
PS
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
XX Claim 1; SEQ ID NO 5961; 1399pp + Sequence Listing; English.
PS
XX
XX The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity/inhibit activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 373 BP; 121 A; 53 C; 74 G; 125 T; 0 other;

Alignment Scores:
Pred. No.: 51.9 Length: 373
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 22 Gaps: 0

US-09-727-892A-99 (1-58) x AA185901 (1-373)

OY 40 GIUTYrAsnLeuPhehrlys 46
|||||
DB 124 GAATATrACTTGTtTACCAAG 144

RESULT 6
ID AAS60503/c
AA60503 standard; cDNA; 509 BP.
AC
XX
XX AAS60503;
DT
XX
XX 29-JAN-2002 (first entry)
DE
XX
XX Human cancer agent-sensitive marker #234.
KW
XX
XX Human; cancer cell marker; TAXOL; cytostatic; tumour; carcinoma;
KW squamous cell carcinoma; sarcoma; fibrosarcoma; leukaemia;
KW lymphocytic leukaemia; lymphoma; plasmacytoma; reticulum cell sarcoma;
KW Hodgkin's disease; glioma; ss.
XX
OS Homo sapiens.

XX
XX
XX WO200179556-A2.
XX
XX 25-OCT-2001.
XX
XX
XX 13-APR-2001; 2001WO-US12132.
XX
XX
XX 14-APR-2000; 2000US-197538P.
XX
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX
XX Lillie J, Brown JL, Bolt A, Van Hufel C;
XX
XX WPI: 2001-602933/68.
XX
XX
XX Novel nucleic acid, used as a marker to determine the effectiveness of
XX using TAXOL to treat cancer cell growth in individuals -
XX
XX
XX Claim 1; Page 232-233; 527pp; English.
XX
XX
XX The invention relates to 1046 novel nucleic acids which are used as
XX markers for determining the sensitivity of a cancer cell to the
XX anticancer agent TAXOL. Cancer cells can be treated with TAXOL when
XX they are shown to express one of the 242 sensitivity markers or the
XX cells are shown not to express one of the 804 resistance markers.
XX The methods can be used to determine the effectiveness of TAXOL
XX in the treatment of cancer cell growth in an individual. The markers
XX can be used as targets in developing anti-cancer agents such as
XX chemotherapeutic compounds. The markers can also be used as targets in
XX developing treatments for cancer, particularly those cancers which
XX display resistance to agents and exhibit expression of the markers. The
XX anticancer agents developed by the novel method can be used to treat
XX cancer. Probes based on the markers can be used to detect transcripts or
XX genomic sequences corresponding to the markers, in the identification of
XX cells or tissues which mis-express the protein. Cancers which may
XX be targeted include carcinoma (e.g. squamous cell carcinoma),
XX sarcoma (e.g. fibrosarcoma) leukaemia (e.g. lymphocytic leukaemia),
XX lymphoma, plasmacytoma, reticulum cell sarcoma, Hodgkin's disease and
XX tumours (e.g. glioma). The present sequence is one of the 1046
XX novel cancer cell markers.
XX
XX
SQ Sequence 509 BP; 150 A; 97 C; 101 G; 161 T; 0 other;

Alignment Scores:
Pred. No.: 68.7 Length: 509
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 22 Gaps: 0

US-09-727-892A-99 (1-58) x AAS60503 (1-509)

OY 6 LysThrValLeuLeuYrCys 12
|||||
DB 486 AAAACAGTCTTGTCTGATTCG 466

RESULT 7
ID AAA01729
AAA01729 standard; cDNA; 788 BP.
AC
XX
XX AAA01729;
DT
XX
XX 19-MAY-2000 (first entry)
DE
XX
XX Human colon cancer cell line polynucleotide sequence SEQ ID NO:1720.
KW
XX
XX Human; colon cancer; tumour; diagnosis; gene expression product;
KW probe; detection; cancerous state; metastasis; identification;
KW breast cancer; oestrogen receptor-positive breast cancer; therapy;
KW oestrogen receptor-negative breast cancer; lung cancer; ss.
XX
OS Homo sapiens.

XX MO9958675-A2.
PN 18-NOV-1999.
XX
PD
XX
PF 13-MAY-1999; 99WO-0510602.
XX
PR 14-MAY-1998; 98US-0085426.
PR 15-MAY-1998; 98US-0085537.
PR 15-MAY-1998; 98US-0085696.
PR 21-OCT-1998; 98US-0105234.
PR 27-OCT-1998; 98US-0105877.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J,
PI Reinhard C, Glese K, Randazzo F, Kennedy GC, Pot D, Kassam A,
PI Lamson G, Drmanac R, Cirvenjakov R, Dickson M, Drmanac S, Labat I,
PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
XX
DR WPI; 2000-126369/11.
XX
XX Polynucleotide library used to determine cancerous states of mammalian
PT cells -
XX
PS Claim 1; Page 629; 1097pp; English.
XX
XX AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
CC libraries constructed from human colon cancer cell lines. The present
CC invention also describes a method of detecting differentially expressed
CC genes correlated with a cancerous state of a mammalian cell, comprising
CC detecting at least one differentially expressed gene product in a test
CC sample derived from a cell suspected of being cancerous, where detection
CC of the differentially expressed gene product is correlated with a
CC cancerous state of the cell from which the test sample was derived.
CC The polynucleotides sequences can be used in a method for detecting
CC differentially expressed genes correlated with a cancerous state of a
CC mammalian cell. The polynucleotides can also be used as probes for
CC detecting and mapping related genes. They can be used in diagnosis and
CC prognosis of diseases and disorders (e.g. identification of
CC pre-metastatic or metastatic cancerous states, stages of cancer, or
CC responsiveness of cancer to therapy). This is particularly for breast
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC negative breast cancer, lung cancer, and colon cancer.
XX
SQ Sequence 788 BP; 260 A; 115 C; 172 G; 222 T; 19 other;
XX
Alignment Scores:
Pred. No.: 102 Length: 788
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 21 Gaps: 0
XX
US-09-727-892a-99 (1-58) x AAA01729 (1-788)
OY 4 LysTyrIysThrValLeuLeu 10
Db 211 AAATATAAGACAGTCTCTA 231
XX
RESULT 8
AAK97705
ID AAK97705 standard; DNA; 970 BP.
XX
AC AAK97705;
XX
DT 13-SEP-1999 (first entry)
XX
DE Extended human secreted protein coding sequence, SEQ ID NO. 270.
XX
KW Secreted protein; human; cytokine; cellular proliferation; cell movement;

KW cellular differentiation; immune system regulator; anti-inflammatory;
KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
KW genetic disease; ss.
XX
OS Homo sapiens.
XX
PN WO9931236-A2.
XX
PD 24-JUN-1999.
XX
PF 17-DEC-1998; 98WO-1B02122.
XX
PR 10-AUG-1998; 98US-0096116.
PR 17-DEC-1997; 97US-0069957.
PR 09-FEB-1998; 98US-0074121.
PR 13-APR-1998; 98US-0081563.
XX
PA (GEST) GENSET.
XX
PI Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX
DR WPI; 1999-385906/32.
DR P-PSDB; AAY36021.
XX
XX New isolated human secreted proteins
PT
XX
PS Claim 1; Page 346-347; 516pp; English.
XX
XX This sequence represents an extended human secreted protein coding
CC sequence of the invention. The secreted proteins can be used in treating
CC or controlling a variety of human conditions. The secreted proteins may
CC act as cytokines or may affect cellular proliferation or differentiation
CC or may act as immune system regulators, haematopoiesis regulators, tissue
CC growth regulators, regulators of reproductive hormones or cell movement
CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
CC tumour inhibition activity. The DNAs can be used in forensic procedures
CC to identify individuals or in diagnostic procedures to identify
CC individuals having genetic diseases resulting from abnormal expression of
CC the genes corresponding to the extended cDNAs. They are also useful for
CC constructing a high resolution map of the human chromosomes. They can
CC also be used for gene therapy to control or treat genetic diseases.
XX
SQ Sequence 970 BP; 267 A; 173 C; 199 G; 323 T; 8 other;
XX
Alignment Scores:
Pred. No.: 123 Length: 970
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 20 Gaps: 0
XX
US-09-727-892a-99 (1-58) x AAK97705 (1-970)
OY 41 TyrAsnLeuPheThrIysLys 47
Db 943 TATTACTTATTTCACAAAAA 963
XX
RESULT 9
AAC45926
ID AAC45926 standard; DNA; 1200 BP.
XX
AC AAC45926;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 48270.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX

OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.

30-JUN-1999; 99US-0141287.
01-JUL-1999; 99US-0141842.
01-JUL-1999; 99US-0142154.
02-JUL-1999; 99US-0142055.
06-JUL-1999; 99US-0142390.
08-JUL-1999; 99US-0142803.
09-JUL-1999; 99US-0142820.
12-JUL-1999; 99US-0142977.
13-JUL-1999; 99US-0143542.
14-JUL-1999; 99US-0143624.
15-JUL-1999; 99US-0144005.
16-JUL-1999; 99US-0144085.
16-JUL-1999; 99US-0144086.
19-JUL-1999; 99US-0144325.
19-JUL-1999; 99US-0144331.
19-JUL-1999; 99US-0144332.
19-JUL-1999; 99US-0144333.
19-JUL-1999; 99US-0144334.
19-JUL-1999; 99US-0144335.
20-JUL-1999; 99US-0144352.
20-JUL-1999; 99US-0144632.
20-JUL-1999; 99US-0144884.
21-JUL-1999; 99US-0144814.
21-JUL-1999; 99US-0145086.
21-JUL-1999; 99US-0145088.
22-JUL-1999; 99US-0145085.
22-JUL-1999; 99US-0145087.
22-JUL-1999; 99US-0145089.
22-JUL-1999; 99US-0145192.
23-JUL-1999; 99US-0145145.
23-JUL-1999; 99US-0145218.
23-JUL-1999; 99US-0145224.
26-JUL-1999; 99US-0145276.
27-JUL-1999; 99US-0145913.
27-JUL-1999; 99US-0145918.
27-JUL-1999; 99US-0145919.
28-JUL-1999; 99US-0145951.
02-AUG-1999; 99US-0146386.
02-AUG-1999; 99US-0146388.
02-AUG-1999; 99US-0146389.
03-AUG-1999; 99US-0147038.
04-AUG-1999; 99US-0147204.
04-AUG-1999; 99US-0147302.
05-AUG-1999; 99US-0147192.
05-AUG-1999; 99US-0147260.
06-AUG-1999; 99US-0147303.
06-AUG-1999; 99US-0147416.
09-AUG-1999; 99US-0147493.
09-AUG-1999; 99US-0147935.
10-AUG-1999; 99US-0148171.
11-AUG-1999; 99US-0148319.
12-AUG-1999; 99US-0148341.
13-AUG-1999; 99US-0148565.
13-AUG-1999; 99US-0148684.
16-AUG-1999; 99US-0149368.
17-AUG-1999; 99US-0149175.
18-AUG-1999; 99US-0149426.
20-AUG-1999; 99US-0149722.
20-AUG-1999; 99US-0149723.
20-AUG-1999; 99US-0149929.
23-AUG-1999; 99US-0149902.
23-AUG-1999; 99US-0149930.
25-AUG-1999; 99US-0150566.
26-AUG-1999; 99US-0150884.
27-AUG-1999; 99US-0151065.
27-AUG-1999; 99US-0151066.
27-AUG-1999; 99US-0151080.
30-AUG-1999; 99US-0151303.
31-AUG-1999; 99US-0151348.
01-SEP-1999; 99US-0151930.
07-SEP-1999; 99US-0152363.
10-SEP-1999; 99US-0153070.
13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159299.
PR 13-OCT-1999; 99US-0159299.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.: 149 Length: 1200
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 21 Gaps: 0

US-09-727-892a-99 (1-58) x AAC45926 (1-1200)

QY 5 TyrLysThrValLeuLeuTyr 11

Db 350 TACAAACGTCGTCCTTTAT 370

RESULT 10

AAC42322
ID AAC42322 standard; DNA; 1754 BP.

AAC42322;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 35111.

XX Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

XX metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN Ep1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.

```

PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143632.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145293.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

```

```

PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Alignment Scores:

```

Pred. No.: 210 Length: 1754
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 21 Gaps: 0

```

US-09-727-892a-99 (1-58) x AAC42322 (1-1754)

Oy 5 TyrlsThrvallentur 11

Db 656 TACAAACGGTCCTTTAT 676

RESULT 11

AAAF21038/c

ID AAAF21038 standard; DNA: 2163 BP.

AAAF21038;

14-MAR-2001 (first entry)

Human low adenosine antisense oligonucleotide related sequence #2605.
 low adenosine antisense oligonucleotide; phosphorothioate; allergy;
 human; airway disorder; bronchoconstriction; lung inflammation;
 surfactant depletion; respiratory; bronchodilator; antiinflammatory;
 immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
 respiratory obstruction; pulmonary obstruction; impeded respiration;
 surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
 respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
 pulmonary hypertension; emphysema; pulmonary transplantation rejection;

KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KM cancer; ss.
XX
OS Homo sapiens.
XX
PN W0200062736-A2.
XX
PD 26-OCT-2000.
XX
PE 24-MAR-2000; 2000WO-US08020.
XX
PR 06-APR-1999; 99US-0127958.
XX
PA (UYEC-) UNIV EAST CAROLINA.
PA (NYCE/) NYCE J W.
PI Nyce JW;
XX
DR WPI; 2000-679539/66.
XX
PT Low adenosine (A) content antisense oligonucleotides which do not
PT trigger adenosine receptors during metabolism, useful e.g. for treating
PT cancers and respiratory obstructions -
XX
PS Disclosure; Page 848; 1592pp; English.
XX
CC The present invention describes low adenosine (A) content antisense
CC oligonucleotides and compositions (I) comprising them. In the antisense
CC oligonucleotides the A is replaced by a 'universal' or alternative base.
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC immunosuppressive, antiallergic, hypotensive and cytostatic activities.
CC The antisense oligonucleotides and (I) can be used to down-regulate the
CC expression and/or activity of target polypeptides associated with
CC lung/respiratory disorders and malignancies, such as stimulating and
CC activating peptide factors and transmitters, transcription factors,
CC immunoglobulins and antibodies, antibody receptors, cytokines and
CC chemokines, endogenously produced specific and non-specific enzymes,
CC binding proteins, adhesion molecules and their receptors, cytokine and
CC chemokine receptors, adenosine receptors, bradykinin receptors, central
CC nervous system (CNS) and peripheral nervous and non-nervous system
CC receptors, CNS and peripheral nervous and non-nervous system peptide
CC transmitters, defensins, growth factors, vasoactive peptides and
CC receptors, binding proteins and malignancy associated proteins. The
CC antisense oligonucleotides may be used in this way to treat disorders
CC including respiratory obstruction (especially pulmonary obstruction
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
CC and/or surfactant hypoproduction which are associated with a disease or
CC condition selected from pulmonary vasoconstriction, inflammation,
CC allergies, asthma, impaired respiration, respiratory distress syndrome
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention.
XX
SQ Sequence 2163 BP; 657 A; 514 C; 482 G; 510 T; 0 other;

Alignment Scores:
Pred. No.: 254 Length: 2163
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 21 Gaps: 0

US-09-727-892a-99 (1-58) x AAF21038 (1-2163)
Oy 16 LysGLYHisPheProHisGln 22
Db 259 AAGGGCAATTTTCACATCAA 239

RESULT 12

AAA34916/C
ID AAA34916 standard; DNA; 2163 BP.
XX
AC AAA34916;
XX
DT 28-JUL-2000 (first entry)
XX
DE Human adenosine receptor related polynucleotide SEQ ID NO:2605.
XX
KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW phosphorothioate; impaired respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW antiallergic; antiallergic; cytostatic; analgesic; impaired airway;
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS Homo sapiens.
XX
PN W0200009525-A2.
XX
PD 24-FEB-2000.
XX
PE 03-AUG-1999; 99WO-US17712.
XX
PR 03-AUG-1998; 98US-0095212.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX
PI Nyce JW;
XX
DR WPI; 2000-205971/18.
XX
PT New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension,
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers -
XX
PS Disclosure; Page 776-777; 1343pp; English.
XX
CC The present invention describes a new composition comprising an
CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
CC inflammation. The ON can have antiinflammatory, antiallergic,
CC antiasthmatic, cytostatic and analgesic activities. The compositions are
CC useful for the treatment of diseases associated with inflammation,
CC impaired airways, including lung disease and diseases whose secondary
CC effects affect the lungs of a subject. They can be used for treating
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
CC asthma, impaired respiration, respiratory distress syndrome, pain, cystic
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC carcinomas, and cancers which may metastasise to the lungs, including
CC breast and prostate cancer. The reduction of the adenosine content of
CC the ONs reduces side effects. The A-containing ONs break down with the
CC release of deoxyadenosine which activates adenosine receptors causing
CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
CC nucleotide sequences given in the sequence listing from the present
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
CC differ from the previously named sequences. SEQ ID NO:11 to 1680
CC (AAA32323 to AAA33992) are specifically claimed ONs from the present
CC invention. N.B. Sequences given in the disclosure of the present
CC invention do not match up with their corresponding SEQ ID NO: sequences
CC given in the sequence listing.
XX
SQ Sequence 2163 BP; 657 A; 514 C; 482 G; 510 T; 0 other;

Alignment Scores:
Pred. No.: 254 Length: 2163
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0

XX	Human secreted protein gene 20 SEQ ID NO:30.
DE	
XX	Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;
KW	antiproliferative; cytosolic; cardiant; vasotropic; cerebroprotective;
KM	nootropic; neuroprotective; antibacterial; vitricide; fungicide;
KW	ophthalmological; vulnerary; autoimmune disease; cardiovascular disorder;
KM	hyperproliferative disorders; cerebrovascular disorder; wound healing;
KW	nervous system disorder; ocular disorder; skin ageing; chemotaxis;
KM	food additive; ss.
KW	
OS	Homo sapiens.
XX	
PN	WO200077026-A1.
XX	
PD	21-DEC-2000.
XX	
PF	01-JUN-2000; 2000NO-US14973.
XX	
PR	11-JUN-1999; 99US-0138630.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
PA	(ROSE/) ROSEN C A.
XX	
PI	Rosen CA, Ruben SM, Komatsoulis GA;
XX	
DR	WPI; 2001-071258/08.
DR	P-PSTDB; AAB75525.
XX	
PT	Nucleic acid molecules encoding human secreted proteins, used in
PT	preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT	Parkinson's diseases and cancers -
PS	
PS	Claim 1; Page 448-449; 542pp; English.
XX	
CC	Human secreted proteins AAB75506 - AAB75554 are encoded by polynucleotide
CC	sequences AAF64176 - AAF64224. The specification includes amino acid
CC	sequences AAB75535 - AAB75606 which represent fragments of the human
CC	secreted proteins, and protein sequences with which they share homology.
CC	The proteins and polynucleotides, their agonists and antagonists have
CC	activities dependent on the tissues and cells in which they are
CC	expressed, examples of these activities include, immunosuppressive;
CC	antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;
CC	vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;
CC	vitricide; fungicide; ophthalmological; and vulnerary. The proteins,
CC	polynucleotides, agonists and antagonists can be used to treat or detect
CC	or diagnose various diseases and disorders including, autoimmune
CC	diseases e.g. Rheumatoid arthritis, hyperproliferative disorders
CC	e.g. neoplasms of the breast or liver, cardiovascular disorders
CC	e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia,
CC	angiogenesis, nervous system disorders e.g. Alzheimer's disease,
CC	infections caused by bacteria, viruses and fungi and ocular disorders
CC	e.g. corneal infection. The polypeptides can also be used to aid wound
CC	healing and epithelial cell proliferation, to prevent skin ageing due to
CC	culture, to maintain organs before transplantation, for supporting cell
CC	turnover of primary tissues, to regenerate tissues and in chemotaxis. The
CC	polypeptides can also be used as a food additive or preservative to
CC	increase or decrease storage capabilities. Included in the invention are
CC	polynucleotide sequences AAF64167 - AAF64175 and peptide AAB75505 which
CC	are used in the isolation, identification and characterisation of the
CC	proteins of the invention.
XX	
SQ	Sequence 2418 BP; 576 A; 500 C; 604 G; 732 T; 6 other;
Alignment Scores:	
Pred. No.:	280
Score:	7.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	12.07%
DB:	22
	Gaps: 0
	Mismatches: 0
	Indels: 0
	Matches: 7
	Length: 2418

(S-09-727-892A-99(1-58) x AAF64195 (1-2418))

QY	33	LysValValTyrSerTyrTyr	39
Db	1384	AAAGTAGTTATCTTATATAT	1404
RESULT	16		
ID	AAF32736		
XX	AAF32736	standard; cDNA; 2418 BP.	
XX	AC		
XX	AAF32736;		
DT	23-MAR-2001	(first entry)	
XX			
DE	Human secreted protein gene 38 SEQ ID NO:48.		
XX			
KW	Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;		
KW	antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;		
KW	neuroprotective; nootropic; neuroprotective; antibacterial; virucide;		
KW	fungicide; ophthalmological; autoimmune disease; rheumatoid arthritis;		
KW	hyperproliferative disorder; neoplasm; cardiovascular disorder;		
KW	cardiac arrest; cerebrovascular disorder; cerebral ischemia; infection;		
KW	angiogenesis; nervous system disorder; Alzheimer's disease; skin aging;		
KW	ocular disorder; corneal infection; wound healing; food additive;		
KW	preservative; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200077255-A1.		
XX			
PD	21-DEC-2000.		
XX			
XX	01-JUN-2000; 2000WO-US14926.		
PE			
PR	11-JUN-1999; 99US-0138628.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
XX			
PI	Rosen CA, Ruden SM, Komatsuolis GA;		
DR	WP1: 2001-025337/03.		
DR	P-PSDB; AAB64459.		
XX			
PT	Isolated nucleic acid molecule encoding a human secreted protein is		
XX	used in preventing, treating or ameliorating a medical condition		
XX			
PS	Claim 1; Page 504; 593pp; English.		
CC	The polynucleotide sequences given in AAF32699 to AAF32747 encode the		
CC	human secreted proteins given in AAB64422 to AAB64470. AAB64471 to		
CC	AAB64548 represent human secreted polypeptide sequences and proteins		
CC	homologous to them, which are given in the exemplification of the present		
CC	invention. Human secreted proteins have activities based on the tissues		
CC	and cells the genes are expressed in. Examples of activities include:		
CC	antiarthritic; immunosuppressive; antirheumatic; antiproliferative;		
CC	cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;		
CC	neuroprotective; antibacterial; virucide; fungicide; and		
CC	ophthalmological. The polynucleotides and polypeptides can be used to		
CC	prevent, treat or ameliorate a medical condition in e.g. humans, mice,		
CC	rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used		
CC	in diagnosing a pathological condition or susceptibility to a		
CC	pathological condition. Disorders which are diagnosed or treated include		
CC	autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative		
CC	disorders e.g. neoplasms of the breast or liver, cardiovascular disorders		
CC	e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia,		
CC	angiogenesis, nervous system disorders e.g. Alzheimer's disease,		
CC	infections caused by bacteria, viruses and fungi and ocular disorders		
CC	e.g. corneal infection. The polypeptides can also be used to aid wound		
CC	healing and epithelial cell proliferation, to prevent skin aging due to		
CC	sunburn, to maintain organs before transplantation, for supporting cell		
CC	culture of primary tissues, to regenerate tissues and in chemotaxis. The		
CC	polypeptides can also be used as a food additive or preservative to		
CC	increase or decrease storage capabilities. AAF32690 to AAF32698 and		
CC	AAB64421 represent sequences used in the exemplification of the present		

CC Invention.
XX
SQ Sequence 2418 BP; 576 A; 500 C; 603 G; 732 T; 7 other;
Alignment Scores:
Pred. No.: 280 Length: 2418
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 22 Gaps: 0
US-09-727-892a-99 (1-58) x AAF32736 (1-2418)
QY 33 LysValValTyrSerTyrTyr 39
|||||
Db 1384 AAGAGTATTATCTTATTAT 1404
RESULT 17
AAD09336/C
ID AAD09336 standard; CDNA: 2433 BP.
XX
AC AAD09336;
XX
DT 10-SEP-2001 (first entry)
XX
DE Rat PRPDE90 cDNA encoding cAMP-specific phosphodiesterase (PDE4) B4.
XX
KW Rat; cyclic AMP-specific phosphodiesterase B4; PDE4B4, antidepressant;
KW memory enhancement; gastrointestinal effect; physiological process; ATP;
KW nausea; intracellular signalling molecule; ss.
XX
OS Rattus norvegicus.
XX
FH Key Location/Qualifiers
FT CDS 262..2241
FT /*tag= a
FT /product= "Rat cyclic AMP (cAMP)-specific
FT phosphodiesterase (PDE4) B4 protein"
FT 262..312
FT /*tag= a
FT /note= "N-terminal region DNA (AAD09337)"
XX
XX WO200144449-A1.
XX
XX 21-JUN-2001.
XX
XX PD 14-DEC-2000; 2000WO-US34045.
XX
XX PR 14-DEC-1999; 99US-0170562.
XX
XX PA (UTAH) UNIV UTAH RES FOUND.
XX
XX PI Bolger G;
XX
XX WPI; 2001-398144/42.
XX
XX P-PSDB; AAE04741.
XX
XX PT Novel rat cyclic AMP-specific phosphodiesterase (PDE4)B protein
XX PT isoforms, PDE4B1 and PDE4B4, useful for identifying inhibitors of cAMP
XX PT phosphodiesterases which have antidepressant and memory enhancement
XX PT effects
XX
XX PS Claim 2; Fig 4; 46pp; English.
XX
XX The present sequence is rat PRPDE90 cDNA encoding cyclic AMP (cAMP)-
XX specific phosphodiesterase (PDE4) B4 protein. The protein molecules are
XX useful for identifying inhibitors of PDE4 which have antidepressant and
XX memory enhancement effects, since the isoforms are expressed in wide
XX range of tissues including various regions of brain. PDE4 inhibitors
XX causes nausea and trigger other gastrointestinal side effects. The
XX phosphodiesterases function in regulation of physiological processes by
XX hydrolysing cAMP, an intracellular signalling molecule derived from ATP.

XX
SQ Sequence 2433 BP; 691 A; 626 C; 591 G; 525 T; 0 other;
Alignment Scores:
Pred. No.: 282 Length: 2433
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 22 Gaps: 0
US-09-727-892a-99 (1-58) x AAD09336 (1-2433)
QY 16 LysGlyHisPheProHisGln 22
|||||
Db 334 AAGGGCATTTCACATCAA 314
RESULT 18
AAO03444
ID AAO03444 standard; DNA: 2497 BP.
XX
AC AAO03444;
XX
DT 17-JUL-1990 (first entry)
XX
DE HindIII S fragment of KS-1 isolate genome.
XX
KW Sheep pox; goat pox; Kenya isolate; S fragment; thymidine kinase; ss.
XX
OS Capripoxvirus, (KS-1 isolate).
XX
FH Key Location/Qualifiers
FT CDS 1..153
FT /*tag= a
FT /label=CF6
FT 110..553
FT /*tag= b
FT /label=CF7
FT 583..1114
FT /*tag= c
FT /label=CF8
FT /note= "corresponds to thymidine kinase gene"
XX
XX CDS 1186..1779
XX
XX FT /*tag= d
XX /label=CF8a
XX 1838..2497
XX FT /*tag= e
XX /label=CF9
XX /*note= "incomplete"
XX
XX GB2222165-A.
XX
XX 28-FEB-1990.
XX
XX PD 20-AUG-1988; 88GB-00196836.
XX
XX PR 20-AUG-1988; 88GB-0019836.
XX
XX PA (ANIM-) INST ANIMAL HEALTH.
XX
XX PI Black DN, Bostox CJ, Gershon PD;
XX
XX WPI; 1990-061358/09.
XX
XX P-PSDB; AAR05130-34.
XX
XX PT Recombinant capripox virus - carrying heterologous gene encoding an
XX PT antigen of pathogen for use in prodn. of vaccines.
XX
XX PS Disclosure; Fig 2; 33pp; English.
XX
XX The sequence is a HindIII S fragment from the capripox viral genome.
XX The capripoxvirus can be used as an alternative to vaccinia as a vector
XX for heterologous genes. The genes can be inserted into the thymidine

CC kinase gene region (CF8) and used for the prodn. of vaccines against a
CC wide variety of diseases in sheep, goats and cattle.
XX
SQ Sequence 2497 BP; 1002 A; 296 C; 391 G; 807 T; 1 other;
Alignment Scores:
Pred. No.: 289 Length: 2497
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 11 Gaps: 0
US-09-727-892a-99 (1-58) x AAF03444 (1-2497)
Qy 42 Asnleuphetrlslystyr 48
Db 1418 AATCTTTTACGAAGAGTAT 1438
RESULT 19
AAF21040/c
ID AAF21040 standard; DNA: 2511 BP.
XX
AC AAF21040;
XX
DT 14-MAR-2001 (first entry)
XX
DE Human low adenosine antisense oligonucleotide related sequence #2607.
XX
KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW human; airway disorder; bronchoconstriction; lung inflammation;
KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW surfactant hypoproduction; pulmonary obstruction; impeded respiration;
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW pulmonary hypertension; emphysema; pain; cystic fibrosis; allergic rhinitis;
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200062736-A2.
XX
PD 26-OCT-2000.
XX
PF 24-MAR-2000; 2000WO-US08020.
XX
PR 06-APR-1999; 99US-0127958.
XX
PA (UYEC-) UNIV EAST CAROLINA.
PA (NYCE/) NYCE J W.
XX
PI NYce JW;
XX
DR WPI; 2000-679539/66.
XX
PT Low adenosine (A) content antisense oligonucleotides which do not
PT trigger adenosine receptors during metabolism, useful e.g. for treating
PT cancers and respiratory obstructions -
XX
PS Disclosure; Page 848-849; 1592pp; English.
XX
XX The present invention describes low adenosine (A) content antisense
CC oligonucleotides and compositions (I) comprising them. In the antisense
CC oligonucleotides the A is replaced by a 'universal' or alternative base.
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC The antisense oligonucleotides and (I) can be used to down-regulate the
CC expression and/or activity of target polypeptides associated with
CC lung/respiratory disorders and malignancies, such as stimulating and
CC activating peptide factors and transmitters, transcription factors,
CC immunoglobulins and antibodies, antibody receptors, cytokines and

CC chemokines, endogenously produced specific and non-specific enzymes,
CC binding proteins, adhesion molecules and their receptors, cytokine and
CC chemokine receptors, adenosine receptors, bradykinin receptors, central
CC nervous system (CNS) and peripheral nervous and non-nervous system
CC receptors, CNS and peripheral nervous and non-nervous system peptide
CC transmitters, defensins, growth factors, vasoactive peptides and
CC receptors, binding proteins and malignancy associated proteins. The
CC antisense oligonucleotides may be used in this way to treat disorders
CC including respiratory obstruction (especially pulmonary obstruction
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
CC and/or surfactant hypoproduction which are associated with a disease or
CC condition selected from pulmonary vasoconstriction, inflammation,
CC allergies, asthma, impeded respiration, respiratory distress syndrome
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention.
XX
SQ Sequence 2511 BP; 745 A; 605 C; 563 G; 598 T; 0 other;
Alignment Scores:
Pred. No.: 290 Length: 2511
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 21 Gaps: 0
US-09-727-892a-99 (1-58) x AAF21040 (1-2511)
Qy 16 LysGLYHisPhePROHisGLn 22
Db 259 AAGGCGCATTTTCCACATCAA 239
RESULT 20
AAA34918/c
ID AAA34918 standard; DNA: 2511 BP.
XX
AC AAA34918;
XX
DT 28-JUL-2000 (first entry)
XX
DE Human adenosine receptor related polynucleotide SEQ ID NO:2607.
XX
KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW phosphorothioate; impaired respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW lung disease; ischemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200009525-A2.
XX
PD 24-FEB-2000.
XX
PF 03-AUG-1999; 99WO-US17712.
XX
PR 03-AUG-1998; 98US-0095212.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX
PI NYce JW;
XX
DR WPI; 2000-205971/18.
XX
PT New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension,

PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
 PT cancers
 XX
 PS Disclosure: Page 777-778; 1343pp; English.
 XX
 CC The present invention describes a new composition comprising an
 CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
 CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
 CC inflammation. The ON can have antiinflammatory, antiallergic,
 CC antistimulant, cytostatic and analgesic activities. The compositions are
 CC useful for the treatment of diseases associated with inflammation.
 CC Impaired airways, including lung disease and diseases whose secondary
 CC effects afflict the lungs of a subject. They can be used for treating
 CC e.g. ischemic conditions, pulmonary vasoconstriction, allergies,
 CC asthma, impaired respiration, respiratory distress syndrome, pain, cystic
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
 CC pulmonary disease (COPD), and cancers such as leukemias, lymphomas,
 CC carcinomas, and cancers which may metastasize to the lungs, including
 CC breast and prostate cancer. The reduction of the adenosine content of
 CC the ONs reduces side effects. The A-containing ONs break down with the
 CC release of deoxyadenosine which activates adenosine receptors causing
 CC bronchoconstriction and inflammation. AAA3313 to AAA3312 represent the
 CC nucleotide sequences given in the sequence listing from the present
 CC invention, which correspond to SEQ ID NO:1 to 185, and then the last
 CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
 CC differ from the previously named sequences. SEQ ID NO:11 to 1680
 CC (AAA2333 to AAA3392) are specifically claimed ONs from the present
 CC invention. N.B. Sequences given in the disclosure of the present
 CC invention do not match up with their corresponding SEQ ID NO: sequences
 CC given in the sequence listing.
 XX
 SO Sequence 2511 BP; 745 A; 605 C; 563 G; 598 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 290 Length: 2511
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 Gaps: 0
 DB: 21
 US-09-727-892A-99 (1-58) x AAA34918 (1-2511)
 QY 16 LysGlyHisPheProHisGln 22
 DB 259 AAGGCCATTTTCACATCAA 239
 RESULT 21
 AAA70146
 ID AAA70146 standard; DNA: 2793 BP.
 XX
 AC AAA70146;
 XX
 DT 07-NOV-2000 (first entry)
 XX
 DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:279.
 XX
 KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
 KW antimalarial; malaria; protozoacide; infection; insecticide; ds.
 XX
 OS Plasmodium falciparum.
 XX
 PN WO200025728-A2.
 PD
 XX 11-MAY-2000.
 XX
 PF 05-NOV-1999; 99WO-US26796.
 XX
 PR 05-NOV-1998; 98US-0107131.
 XX
 PA (HOFF/) HOFFMAN S.
 PA (CARU/) CARUCCI D.
 PA (GARD/) GARDNER M.

(VENT/) VENTER J C.
 Hoffman S, Carucci D, Gardner M, Venter JC;
 WPI; 2000-365347/31.
 Proteins encoded by chromosome 2 of the human malarial parasite,
 Plasmodium falciparum, useful as antimalarial vaccines and in the
 diagnosis of P.falciparum infection
 Disclosure: Page 488; 577pp; English.
 XX
 CC The present invention describes proteins and their fragments (I) encoded
 CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
 CC Also described are: (I) nucleotide sequences (II) encoding (I); and (2)
 CC vaccines against P. falciparum infection comprising (I) or (II).
 CC (I) and (II) are useful for the development of vaccines against
 CC P. falciparum infection. (I) and polyclonal antisera or a monoclonal
 CC antibody raised to immunogens comprising the sequences of (I), are
 CC useful in the detection of infection with P. falciparum. Furthermore,
 CC (I) (especially when they are rifins or secreted or membrane proteins)
 CC can aid the identification of drugs to treat or prevent P. falciparum
 CC infection, or they can be used to identify drug resistance in
 CC P. falciparum. Sequencing of the plasmodium chromosome 2 and the
 CC subsequent identification of proteins encoded by it will help to expand
 CC our understanding of parasite biology, a process hampered by the
 CC complexity of the parasitic lifecycle, and provide new targets for
 CC vaccine and drug development. Parasite resistance to drugs and mosquito
 CC resistance to insecticides have led to a resurgence of malaria in many
 CC parts of the world, and there is a pressing need for vaccines and new
 CC drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352 represent nucleotide
 CC and protein sequences given in the present invention, but which are not
 CC specifically mentioned within the specification.
 XX
 SO Sequence 2793 BP; 1184 A; 355 C; 440 G; 814 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 319 Length: 2793
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 Gaps: 0
 DB: 21
 US-09-727-892A-99 (1-58) x AAA70146 (1-2793)
 QY 1 MetGluArgLysTyrLysThr 7
 DB 2069 ATGCAAGAAAGATCAAAACT 2089
 RESULT 22
 AAD09338/C
 ID AAD09338 standard; cDNA; 3022 BP.
 XX
 AC AAD09338;
 XX
 DT 10-SEP-2001 (first entry)
 XX
 DE Rat PRPDB89 cDNA encoding cAMP-specific phosphodiesterase (PDE4) B1.
 XX
 KW Rat; cyclic AMP-specific phosphodiesterase B1; PDE4B1; antidepressant;
 KW memory enhancement; gastrointestinal effect; physiological process; ATP;
 KW nausea; intracellular signalling molecule; ss.
 XX
 OS Rattus norvegicus.
 XX
 PN WO200144449-A1.
 Key Location/Qualifiers
 CDS 325..2535
 FT /tag= a
 FT /product= "Rat cyclic AMP (cAMP)-specific
 FT phosphodiesterase (PDE4) B1 protein"
 FT
 XX
 XX

```
XX 21-JUN-2001.
PD 14-DEC-2000; 2000WO-US34045.
XX
PF 14-DEC-1999; 99US-0170562.
XX
PR 14-DEC-1999; 99US-0170562.
XX
PA (UTAH ) UNIV UTAH RES FOUND.
XX
PI Bolger G;
XX
DR WPI; 2001-398144/42.
DR P-PSDB; AAE04743.
XX
XX Novel rat cyclic AMP-specific phosphodiesterase (PDE)4B protein
PT isoforms, PDE4B1 and PDE4B4, useful for identifying inhibitors of cAMP
PT phosphodiesterases which have antidepressant and memory enhancement
PT effects -
XX
XX Claim 14; Fig 5; 46pp; English.
XX
CC The present sequence is rat PRPDE89 cDNA encoding cyclic AMP (cAMP)-
CC specific phosphodiesterase (PDE4) B1 protein. The protein molecules are
CC useful for identifying inhibitors of PDE4 which have antidepressant and
CC memory enhancement effects, since the isoforms are expressed in wide
CC range of tissues including various regions of brain. PDE4 inhibitors
CC causes nausea and trigger other gastrointestinal side effects. The
CC phosphodiesterases function in regulation of physiological processes by
CC hydrolysing cAMP, an intracellular signalling molecule derived from ATP.
XX
SQ Sequence 3022 BP; 827 A; 761 C; 771 G; 663 T; 0 other;

Alignment Scores:
Pred. No.: 343 Length: 3022
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: Gaps: 0

US-09-727-892a-99 (1-58) x AAD09338 (1-3022)

OY 16 LysGlyHisPheProHisGln 22
DB 628 AAGGCCATTTTCACATCAA 608

RESULT 23
AAD09340/C
ID AAD09340 standard; cDNA; 3133 BP.
XX
AC AAD09340;
XX
DT 10-SEP-2001 (first entry)
XX
DE Rat PRPDE74 cDNA encoding cAMP-specific phosphodiesterase (PDE4) B3.
XX
KW Rat; cyclic AMP-specific phosphodiesterase B3; PDE4B3; antidepressant;
KW memory enhancement; gastrointestinal effect; physiological process; ATP;
KW nausea; intracellular signalling molecule; ss.
XX
OS Rattus norvegicus.
XX
XX Key location/Qualifiers
XX CDS 512..2677
XX /tag= a
XX /product= "Rat cyclic AMP (cAMP)-specific
XX phosphodiesterase (PDE4) B3 protein"
XX
XX MO20014449-A1.
XX
XX 21-JUN-2001.
XX
XX 14-DEC-2000; 2000WO-US34045.
XX
```

```
XX 14-DEC-1999; 99US-0170562.
XX
XX (UTAH ) UNIV UTAH RES FOUND.
XX
XX Bolger G;
XX
DR WPI; 2001-398144/42.
DR P-PSDB; AAE04746.
XX
XX Novel rat cyclic AMP-specific phosphodiesterase (PDE)4B protein
PT isoforms, PDE4B1 and PDE4B4, useful for identifying inhibitors of cAMP
PT phosphodiesterases which have antidepressant and memory enhancement
PT effects -
XX
XX Example 1; Page 42-44; 46pp; English.
XX
CC The present sequence is rat PRPDE74 cDNA encoding cyclic AMP (cAMP)-
CC specific phosphodiesterase (PDE4) B3 protein. The protein molecules are
CC useful for identifying inhibitors of PDE4 which have antidepressant and
CC memory enhancement effects, since the isoforms are expressed in wide
CC range of tissues including various regions of brain. PDE4 inhibitors
CC causes nausea and trigger other gastrointestinal side effects. The
CC phosphodiesterases function in regulation of physiological processes by
CC hydrolysing cAMP, an intracellular signalling molecule derived from ATP.
XX
SQ Sequence 3133 BP; 878 A; 787 C; 739 G; 729 T; 0 other;

Alignment Scores:
Pred. No.: 354 Length: 3133
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: Gaps: 0

US-09-727-892a-99 (1-58) x AAD09340 (1-3133)

OY 16 LysGlyHisPheProHisGln 22
DB 770 AAGGCCATTTTCACATCAA 750

RESULT 24
AAQ14631/C
ID AAQ14631 standard; DNA; 3186 BP.
XX
AC AAQ14631;
XX
DT 30-JAN-1992 (first entry)
XX
DE Plasmid pTM72 insert encoding a human cAMP phosphodiesterase.
XX
KW Cyclic nucleotide; PDE; glioblastoma cell; rolipram-sensitive; ss.
XX
OS Homo sapiens.
XX
XX Key location/Qualifiers
XX CDS 139..2348
XX /tag= a
XX
XX MO9116457-A.
XX
XX 31-OCT-1991.
XX
XX 19-APR-1991; 91WO-US02714.
XX
XX 20-APR-1990; 90US-0511715.
XX
XX (COLD-) COLD SPRING HARBOR.
XX
XX Wigler MH, Colicelli J;
XX
XX WPI; 1991-339841/46.
XX
```

DR P-PSDB; AAR14843.
XX Complementary screening for genes and prods. - e.g. RAS protein
PT and CAMP, that modify, complement or suppress genetic defect and
PT correct associated phenotypic alteration
XX
PS Claim 15; Page 91; 169pp; English.
XX
CC Plasmid pTM72 contains a human glioblastoma cDNA which encodes a
CC rolipram-sensitive CAMP PDE. The sequence is very closely related
CC to, but distinct from, the rat PRATPD cDNA insert (see A014624).
CC N.B. This sequence is SEQ ID NO. 22 in the specification but is
CC referred to as SEQ ID NO. 23 in the text, e.g. on page 32.
XX
SQ Sequence 3186 BP; 927 A; 743 C; 734 G; 782 T; 0 other;
XX
Alignment Scores:
Pred. No.: 360 Length: 3186
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 12 Gaps: 0
US-09-727-892a-99 (1-58) x A014631 (1-3186)
QY 16 LysGlyHisPheProHisGln 22
Db 442 AAGGGCCATTTCACATCAA 422
RESULT 25
AAT34375/c
ID AAT34375 standard; cDNA; 3186 BP.
XX
AC AAT34375;
XX
DT 09-OCT-1996 (first entry)
XX
DE Plasmid pTM72 (ATCC 68602) insert.
XX
KW Human: glioblastoma cell; heat shock sensitivity; phosphodiesterase;
KW deficient yeast strain 10DAB; pTM22; rat pPD phosphodiesterase; pde1-;
KW bovine Ca2+/calmodulin dependent CAMP phosphodiesterase; heart; plasmid;
KW RAS2(val19); pde2-; pTM3; pTM72; PRATPD; pJC99; rolipram sensitive; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 139..2349
FT /*tag= a
FT /product= CAMP phosphodiesterase
XX
PN US5527896-A.
XX
PD 18-JUN-1996.
XX
PE 20-APR-1990; 90US-0511715.
XX
PR 19-APR-1991; 91US-0688352.
PR 20-APR-1990; 90US-0511715.
XX
PA (COLD-) COLD SPRING HARBOR LAB.
XX
PI Colicelli JT, Wigler MH;
XX
DR WPI; 1996-289902/30.
DR P-PSDB; AAW00093.
XX
PT DNA mols. isolated from human glioblastoma cells - encode
PT RAS-related or cyclic nucleotide phosphodiesterase proteins
XX
PS Claim 4; Column 81-88; 101pp; English.
XX

EC The sequences given in AAT34374-76 represent plasmid fragments which
EC contain human glioblastoma cell cDNA inserts which are capable of
EC correcting the heat shock sensitivity of the phosphodiesterase
EC deficient yeast strain 10DAB. Several cDNA's were isolated and
EC sequenced. pTM22 encodes a novel human gene. From computer analysis,
EC pTM22 putatively encodes a protein homologous to various CAMP
EC phosphodiesterases, such as the bovine Ca2+/calmodulin dependent CAMP
EC phosphodiesterase and the rat pPD phosphodiesterase. Sequences related
EC to pTM22 were found to be expressed in human heart. Plasmid pTM22 was
EC unable to correct the heat shock sensitivity of RAS2(val19) yeast
EC strains. It thus appears that the pde1- and pde2- yeast strain 10DAB
EC is more sensitive to phenotypic reversion by mammalian CAMP
EC phosphodiesterase clones than is the RAS2(val19) yeast strain. The
EC inserts in the plasmids pTM3 and pTM72 were also characterised. These
EC two different CAMP phosphodiesterase cDNA's were found to be closely
EC related to, but distinct from, the PRATPD insert and the pJC99 insert.
EC Biochemical analysis of cell lysates has established that the cDNA's of
EC pTM3 and pTM72, pJC44x and PRATPD encode rolipram sensitive CAMP
EC phosphodiesterases.
XX
SQ Sequence 3186 BP; 927 A; 744 C; 734 G; 781 T; 0 other;
XX
Alignment Scores:
Pred. No.: 360 Length: 3186
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 17 Gaps: 0
US-09-727-892a-99 (1-58) x AAT34375 (1-3186)
QY 16 LysGlyHisPheProHisGln 22
Db 442 AAGGGCCATTTCACATCAA 422
RESULT 26
AAZ32242/c
ID AAZ32242 standard; cDNA; 3186 BP.
XX
AC AAZ32242;
XX
DT 19-JAN-2000 (first entry)
XX
DE Human glioblastoma cell CAMP phosphodiesterase pTM72 encoding cDNA.
XX
KW Phosphodiesterase; dunce-like phosphodiesterase; PDE; pPD; CAMP;
KW RAS-related protein; immunoreactive; detection; genetic defect;
KW bronchodilation; increased myocardial contractility;
KW anti-inflammation; ss.
XX
OS Homo sapiens.
XX
PN US5977305-A.
XX
PD 02-NOV-1999.
XX
PE 07-JUN-1995; 95US-0474379.
XX
PR 01-MAR-1994; 94US-0206188.
PR 20-APR-1990; 90US-0511715.
PR 19-APR-1991; 91US-0688352.
XX
PA (COLD-) COLD SPRING HARBOR LAB.
XX
PI Colicelli JT, Wigler MH;
XX
DR WPI; 1999-619709/53.
DR P-PSDB; AAY49810.
XX
PT New isolated RAS-related polypeptides and mammalian cyclic nucleotide
PT phosphodiesterases, used for screening for agents which can modify
PT complement or suppress genetic defects -
XX

CC participates and are also used as hybridisation probes. The present
CC invention also describes methods for detecting mammalian genes encoding
CC proteins which can function in microorganisms, particularly yeast, to
CC modify, complement, or suppress a genetic defect associated with an
CC identifiable phenotypic alteration or characteristic in the
CC microorganism. AAA88162 to AAA88218 and AAB29614 to AAB20640 represent
CC sequences used in the exemplification of the present invention.

XX SQ Sequence 3233 BP; 930 A; 745 C; 735 G; 784 T; 2 U; 37 other;

Alignment Scores:

Pred. No.:	365	Length:	3233
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.07%	Indels:	0
DB:	21	Gaps:	0

US-09-727-892a-99 (1-58) x AAA88177 (1-3233)

OY 16 LysGlyHisPheProHisGln 22

DB 442 AAGGCCATTTCACATCAA 422

RESULT 29

ABLO9204
ID ABL09204 standard; cDNA; 5080 BP.

XX ABL09204;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 22094.

XX Drosophila: developmental biology; cell signalling; insecticide;

KM pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR P-PSDB; ABB65101.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX Claim 1; SEQ ID NO 22094; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 5080 BP; 1504 A; 1200 C; 1171 G; 1205 T; 0 other;

Alignment Scores:

Pred. No.:	548	Length:	5080
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.07%	Indels:	0
DB:	23	Gaps:	0

US-09-727-892a-99 (1-58) x ABL09204 (1-5080)

OY 10 LeuTYrCysAspGluIleLys 16

DB 2814 CTTTATTCGTGATGAATCAAA 2834

RESULT 30

ABLI4026/C
ID ABLI4026 standard; cDNA; 5536 BP.

XX ABLI4026;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 36560.

XX Drosophila: developmental biology; cell signalling; insecticide;

KM pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR P-PSDB; ABB69923.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX Claim 1; SEQ ID NO 36560; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

Alignment Scores:

Pred. No.:	592	Length:	5536
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.07%	Indels:	0


```
DB: 23 Gaps: 0
US-09-727-892a-99 (1-58) x ABL14026 (1-5536)
OY 40 GATGTAAATTTGTTTACCAA 1045
DB 1065 GATGTAAATTTGTTTACCAA 1045
RESULT 31
ABL32591/C
ID ABL32591 standard; DNA; 6195 BP.
XX
XX ABL32591;
XX
XX 26-MAR-2002 (first entry)
XX
XX Human immune system associated gene SEQ ID NO: 564.
XX
XX Human; immune system disease; cytosine methylation; antiasthmatic;
XX antiarteriosclerotic; antihaemic; cytosolic; nootropic;
XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
XX antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
XX antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
XX gene; ds.
XX
XX Homo sapiens.
XX
XX WO200200928-A2.
XX
XX 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP07537.
XX
XX 30-JUN-2000; 2000DE-1032529.
XX
XX 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful
XX for diagnosis and treatment of diseases associated with abnormal
XX cytosine methylation -
XX
XX Claim 1; SEQ ID NO 564; 32pp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
XX can be used in the diagnosis and treatment of immune system disorders,
XX including eye diseases such as retinopathy, neovascular glaucoma and
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX diseases. The present sequence is a gene of the invention.
XX
XX SQ Sequence 6195 BP; 1636 A; 57 C; 1209 G; 3293 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 656 Length: 6195
XX Score: 7.00 Matches: 7
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 12.07% Indels: 0
XX DB: 24 Gaps: 0
XX
XX US-09-727-892a-99 (1-58) x ABL32591 (1-6195)
XX
XX OY 42 AsnLeuPheThrLysLysTyr 48
XX |||||||||||||||||||
DB: 4219 AATCTTTTCACCTAAATAATAT 4199
RESULT 32
ABL32866
ID ABL32866 standard; DNA; 6219 BP.
XX
XX ABL32866;
XX
XX 26-MAR-2002 (first entry)
XX
XX Human immune system associated gene SEQ ID NO: 839.
XX
XX Human; immune system disease; cytosine methylation; antiasthmatic;
XX antiarteriosclerotic; antihaemic; cytosolic; nootropic;
XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
XX antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
XX antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
XX gene; ds.
XX
XX Homo sapiens.
XX
XX WO200200928-A2.
XX
XX 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP07537.
XX
XX 30-JUN-2000; 2000DE-1032529.
XX
XX 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful
XX for diagnosis and treatment of diseases associated with abnormal
XX cytosine methylation -
XX
XX Claim 1; SEQ ID NO 839; 32pp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
XX can be used in the diagnosis and treatment of immune system disorders,
XX including eye diseases such as retinopathy, neovascular glaucoma and
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX diseases. The present sequence is a gene of the invention.
XX
XX SQ Sequence 6219 BP; 1714 A; 141 C; 1278 G; 3086 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 658 Length: 6219
XX Score: 7.00 Matches: 7
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 12.07% Indels: 0
XX DB: 24 Gaps: 0
XX
XX US-09-727-892a-99 (1-58) x ABL32866 (1-6219)
XX
XX OY 52 IleGluTyrIleLysGluIle 58
XX |||||||||||||||||||
DB 2661 AFTGAATATATTAAAGCAAT 2681
RESULT 33
AAS63324
ID AAS63324 standard; DNA; 6219 BP.
XX
```

AC AAS63324;
 XX
 DT 29-JAN-2002 (first entry)
 XX
 DE Chemically pretreated metabolism associated gene #19.
 XX
 KW Human; cytostatic; anti-tumour; metabolism; metabolic disease; liver;
 KW solid tumour; cancer; cytosine methylation; epigenetic; eye; kidney;
 KW single nucleotide polymorphism detection; SNP; stool; urine; lung;
 KW cerebral-spinal fluid; intestine; brain; prostate; breast;
 KW DUSP2; EPHX2; GDSH; SHMT2; SLC7A2; SLC7A4; TYMS; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200176451-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001; 2001WO-EP04016.
 XX
 PR 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIC-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI; 2002-010834/01.
 XX
 PT New nucleic acid, useful for diagnosis and therapy of metabolic
 PT disease, solid tumour and cancers, comprises segment of chemically
 PT modified genomic sequences of genes associated with metabolism
 XX
 PS Claim 1; Page 64-65; 143pp; English.
 XX
 CC The invention relates to a nucleic acid (I) comprising a sequence at
 CC least 18 bases of a segment of the chemically pretreated DNA of genes
 CC associated with metabolism such as DUSP2 (NM_004418), EPHX2 (NM_001979),
 CC ODP2 (NM_000320), GDSH (NM_000199), SHMT2 (NM_005412), SLC7A2
 CC (NM_003046), SLC7A4 (NM_004173) and TYMS (NM_001071) (all
 CC undefined). (I) are useful for diagnosis and therapy of metabolic
 CC disease, solid tumours and cancers; as primer oligonucleotides for the
 CC amplification of DNA sequences, for detecting the cytosine methylation
 CC state and/or single nucleotide polymorphisms (SNPs) in a chemically
 CC treated DNA of genes associated with metabolism. An array of (I) is
 CC useful for ascertaining genetic and/or epigenetic parameters for the
 CC diagnosis and/or therapy of existing diseases or the predisposition to
 CC specific diseases by analysing cytosine methylations. The method involves
 CC chemically treating genomic DNA sample by a solution of bisulphite,
 CC hydrogen sulphite or disulphite such that cytosine bases which are
 CC unmethylated at the 5th-position are converted to uracil or another base
 CC which is dissimilar to cytosine in terms of hybridisation behaviour and
 CC amplifying fragments of the chemically pretreated genomic DNA. The
 CC genomic DNA is from cells or cellular components which contain DNA,
 CC sources of DNA comprising, for e.g. cell lines, biopsies, blood, sputum,
 CC stool, urine, cerebral-spinal fluid, tissue embedded in paraffin such as
 CC tissue from eye, intestine, kidney, brain, heart, prostate, lung, breast
 CC or liver, histologic object slides and their combinations. Genetic
 CC parameters are mutations, in particular insertions, deletions, point
 CC mutations, inversions and polymorphisms of genes associated with
 CC metabolism and sequences further required for their regulation.
 CC Epigenetic parameters are in particular cytosine methylations and
 CC further chemical modifications of DNA bases of genes associated with
 CC metabolism. Further epigenetic parameters include for e.g. the
 CC acetylation of histones which correlates with DNA methylation.
 CC AAS63306-AAS63373 represent chemically pretreated metabolism associated
 CC genes, and related primers of the invention.
 XX
 SQ Sequence 6219 BP; 1714 A; 141 C; 1278 G; 3086 T; 0 other;

Alignment Scores:

Pred. No.: 658 Length: 6219
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 DB: 24 Gaps: 0
 US-09-727-892a-99 (1-58) x AAS63324 (1-6219)
 QY 52 IleguTYrlleYsguile 58
 ||||||||||||||||
 DB 2661 ATTGAATATATTAGCAATT 2681
 RESULT 34
 ABL15218/c
 ID ABL15218 standard; cDNA: 6447 BP.
 XX
 AC ABL15218;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 40136.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW,
 XX
 DR WPI; 2001-656860/75.
 XX
 PT P-PSDB; ABB71115.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Claim 1; SEQ ID NO 40136; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABB57737-ABB72072).
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/publshd_pcl_sequences.
 XX
 SQ Sequence 6447 BP; 1710 A; 1492 C; 1537 G; 1708 T; 0 other;

Alignment Scores:

Pred. No.: 680 Length: 6447
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 DB: 23 Gaps: 0

US-09-727-892a-99 (1-58) x ABL15218 (1-6447)

OY 40 GluTyrAsnLeuPheThrLys 46
|||||
Db 5650 GAATATAAATTTGTTATCCCAA 5630
RESULT 35
ABL22283
ID ABL22283 standard; DNA; 6519 BP.
XX
AC ABL22283;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 18332.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
XX OS
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX PD
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX PI
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI: 2001-656860/75.
XX DR
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Claim 1; SEQ ID NO 18332; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB161716-AB130511), expressed DNA
CC sequences (AB101840-AB161715) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 6519 BP; 2177 A; 1452 C; 1333 G; 1557 T; 0 other;
XX SQ
Alignment Scores:
Pred. No.: 686 Length: 6519
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 23 Gaps: 0
US-09-727-892a-99 (1-58) x ABL22283 (1-6519)
OY 38 TyrTyrGluTyrAsnLeuPhe 44
|||||
Db 4848 TATTACGAGTACAACTATTC 4868
RESULT 36
AAK85381
ID AAK85381 standard; DNA; 6581 BP.
XX

AC AAK85381;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:40193.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX PD
XX PF 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180628.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-APR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 07-JUN-2000; 2000US-0209467.
XX PR 28-JUN-2000; 2000US-0214886.
XX PR 30-JUN-2000; 2000US-0215135.
XX PR 07-JUL-2000; 2000US-0216647.
XX PR 07-JUL-2000; 2000US-0216880.
XX PR 11-JUL-2000; 2000US-0217487.
XX PR 11-JUL-2000; 2000US-0217496.
XX PR 14-JUL-2000; 2000US-0218290.
XX PR 26-JUL-2000; 2000US-0220963.
XX PR 26-JUL-2000; 2000US-0220964.
XX PR 14-AUG-2000; 2000US-0224518.
XX PR 14-AUG-2000; 2000US-0224519.
XX PR 14-AUG-2000; 2000US-0225213.
XX PR 14-AUG-2000; 2000US-0225214.
XX PR 14-AUG-2000; 2000US-0225266.
XX PR 14-AUG-2000; 2000US-0225267.
XX PR 14-AUG-2000; 2000US-0225268.
XX PR 14-AUG-2000; 2000US-0225270.
XX PR 14-AUG-2000; 2000US-0225447.
XX PR 14-AUG-2000; 2000US-0225757.
XX PR 14-AUG-2000; 2000US-0225758.
XX PR 14-AUG-2000; 2000US-0225759.
XX PR 18-AUG-2000; 2000US-0226279.
XX PR 22-AUG-2000; 2000US-0226681.
XX PR 22-AUG-2000; 2000US-0226868.
XX PR 22-AUG-2000; 2000US-0227182.
XX PR 23-AUG-2000; 2000US-0227009.
XX PR 30-AUG-2000; 2000US-0228924.
XX PR 01-SEP-2000; 2000US-0229287.
XX PR 01-SEP-2000; 2000US-0229343.
XX PR 01-SEP-2000; 2000US-0229344.
XX PR 01-SEP-2000; 2000US-0229345.
XX PR 05-SEP-2000; 2000US-0229509.
XX PR 05-SEP-2000; 2000US-0229513.
XX PR 06-SEP-2000; 2000US-0230437.
XX PR 06-SEP-2000; 2000US-0230438.
XX PR 08-SEP-2000; 2000US-0231242.
XX PR 08-SEP-2000; 2000US-0231243.
XX PR 08-SEP-2000; 2000US-0231244.
XX PR 08-SEP-2000; 2000US-0231413.
XX PR 08-SEP-2000; 2000US-0231414.
XX PR 08-SEP-2000; 2000US-0232080.
XX PR 08-SEP-2000; 2000US-0232081.
XX PR 12-SEP-2000; 2000US-0231968.
XX PR 14-SEP-2000; 2000US-0232397.
XX PR 14-SEP-2000; 2000US-0232398.
XX PR 14-SEP-2000; 2000US-0232399.
XX PR 14-SEP-2000; 2000US-0232400.

PF 17-JAN-2001; 2001MO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241222.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246539.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
PA
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,

PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 40195; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 6581 BP; 2160 A; 1264 C; 1239 G; 1918 T; 0 other;

Alignment Scores:
Pred. No.: 692 Length: 6581
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 22 Gaps: 0

US-09-727-892a-99 (1-58) x AAK85383 (1-6581)

QY 46 LysLysTyrAlaTyrIleIle 52
DB 3881 AAAAAGTATGCTTATATTATA 3901

RESULT 38
ABL22879
ID ABL22879 standard; DNA: 7275 BP.
XX
AC ABL22879;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 20110.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PE 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX

Claim 1; SEQ ID NO 20110; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB101840-AB116175), expressed DNA
XX sequences (AB101840-AB116175) and the encoded proteins
XX (AB157737-AB172072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 7275 BP; 1980 A; 1650 C; 1792 G; 1853 T; 0 other;

Alignment Scores:
Pred. No.: 758 Length: 7275
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 23 Gaps: 0

US-09-727-892a-99 (1-58) x ABL22879 (1-7275)

Y 27 GluAspLeuTyrAspAlaLys 33
DB 3616 GAAGACCTCTATGACGCCAA 3636

RESULT 39
AAK84675/C
ID AAK84675 standard; DNA: 9001 BP.
XX
AC AAK84675;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:39487.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PE 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
XX
PR 04-FEB-2000; 2000US-0180628.
XX
PR 24-FEB-2000; 2000US-0184664.
XX
PR 02-MAR-2000; 2000US-0186350.
XX
PR 16-MAR-2000; 2000US-0189874.
XX
PR 17-MAR-2000; 2000US-0190076.
XX
PR 18-APR-2000; 2000US-0198123.
XX
PR 19-MAY-2000; 2000US-0205515.
XX
PR 07-JUN-2000; 2000US-0209467.
XX
PR 28-JUN-2000; 2000US-0214886.
XX
PR 30-JUN-2000; 2000US-0215135.
XX
PR 07-JUL-2000; 2000US-0216647.
XX
PR 07-JUL-2000; 2000US-0216880.
XX
PR 11-JUL-2000; 2000US-0217487.
XX
PR 11-JUL-2000; 2000US-0217496.
XX
PR 14-JUL-2000; 2000US-0218290.
XX
PR 26-JUL-2000; 2000US-0220963.
XX
PR 26-JUL-2000; 2000US-0220964.
XX
PR 14-AUG-2000; 2000US-0224518.
XX
PR 14-AUG-2000; 2000US-0224519.
XX
PR 14-AUG-2000; 2000US-0225213.
XX
PR 14-AUG-2000; 2000US-0225214.

[illegible]

PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246529.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
XX Rosen CA, Barash SC, Ruben SM.
XX
XX WPI: 2001-483426/52.
XX
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
XX
XX Disclosure; SEQ ID NO 39487; 3071pp + Sequence Listing; English.
XX
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytosstatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87594 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX represent sequences used in the exemplification of the present invention.
XX

Alignment Scores:

Pred. No.:	919	Length:	9001
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.07%	Indels:	0
DB:	22	Gaps:	0

US-09-727-892a-99 (1-58) x AAK84675 (1-9001)

QY 6 LysThrValLeuLeuTyrCys 12
|||||

Db 8025 AAAACAGCTTGCTATATGTC 8005

RESULT 40
ABL22878

ID ABL22878 standard; DNA; 11424 BP.

XX
AC ABL22878;

XX
DT 26-MAR-2002 (first entry)

XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 20107.

XX
KW Drosophila; developmental biology; cell signalling; insecticide;

XX
KW pharmaceutical; gene; ds.

OS Drosophila melanogaster.

XX
PN WO200171042-A2.

XX
PD 27-SEP-2001.

XX
PF 23-MAR-2001; 2001WO-US09231.

XX
PR 23-MAR-2000; 2000US-191637P.

XX
PR 11-JUL-2000; 2000US-0614150.

XX
PA (PEKE) PE CORP NY.

XX
PI Venter JC, Adams M, Li PWD, Myers EW;

XX
DR WPI; 2001-656860/75.

XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

XX
PS Claim 1; SEQ ID NO 20107; 21pp + Sequence Listing; English.

XX
CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL10511), expressed DNA

CC sequences (ABR101840-ABR16175) and the encoded proteins

CC (ABR57737-ABR72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 11424 BP; 3193 A; 2399 C; 2656 G; 3176 T; 0 other;

Alignment Scores:

Pred. No.:	1.14e+03	Length:	11424
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.07%	Indels:	0
DB:	23	Gaps:	0

(S-09-727-892a-99 (1-58) x ABL22878 (1-11424))

(7 27 GLuAspleuTyrAspAlaLys 33
|||||

(2 5979 GAAGACCTCTATGAGCCAA 5999

RESULT 41
ABA14848/c

ABA14848 standard; DNA; 11641 BP.

ABA14848;

23-JAN-2002 (first entry)

Human nervous system related polynucleotide SEQ ID NO 7179.

Human: nootropic; neuroprotective; cytostatic; dermatological; vitruicide;
Immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerrary;
antiparkinsonian; antistickling; antianaemic; antirheumatic; cancer;
antipneumatic; hepatotropic; cerebroprotective; antiinflammatory;
antiallergic; antidiabetic; antileuc; anticonvulsant; antifungal;
antiparasitic; cardiac; immune disorder; cardiovascular disorder;
neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.

Homo sapiens.

WO200159063-A2.

16-AUG-2001.

17-JAN-2001; 2001WO-US01334.

31-JAN-2000; 2000US-0179065.

04-FEB-2000; 2000US-0180628.

24-FEB-2000; 2000US-0184664.

02-MAR-2000; 2000US-0186350.

16-MAR-2000; 2000US-0189874.

17-MAR-2000; 2000US-0190076.

18-APR-2000; 2000US-0198123.

19-MAY-2000; 2000US-0205515.

07-JUN-2000; 2000US-0209467.

28-JUN-2000; 2000US-0214886.

30-JUN-2000; 2000US-0215135.

07-JUL-2000; 2000US-0216647.

11-JUL-2000; 2000US-0217487.

11-JUL-2000; 2000US-0217496.

14-JUL-2000; 2000US-0218290.

26-JUL-2000; 2000US-0220963.

26-JUL-2000; 2000US-0220964.

14-AUG-2000; 2000US-0224518.

14-AUG-2000; 2000US-0224519.

14-AUG-2000; 2000US-0225213.

14-AUG-2000; 2000US-0225214.

14-AUG-2000; 2000US-0225266.

14-AUG-2000; 2000US-0225267.

14-AUG-2000; 2000US-0225268.

14-AUG-2000; 2000US-0225270.

14-AUG-2000; 2000US-0225447.

14-AUG-2000; 2000US-0225757.

14-AUG-2000; 2000US-0225758.

14-AUG-2000; 2000US-0225759.

18-AUG-2000; 2000US-0226279.

22-AUG-2000; 2000US-0226681.

22-AUG-2000; 2000US-0226868.

22-AUG-2000; 2000US-0227182.

30-AUG-2000; 2000US-0227009.

30-AUG-2000; 2000US-0228924.

01-SEP-2000; 2000US-0229287.

01-SEP-2000; 2000US-0229343.

01-SEP-2000; 2000US-0229344.

01-SEP-2000; 2000US-0229345.

05-SEP-2000; 2000US-0229509.

[illegible]

PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251130.
PR 05-DEC-2000; 2000US-0251388.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251889.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-541565/60.
XX
XX
XX Nucleic acids encoding 3324 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX
XX
PS Disclosure; SEQ ID NO 7179; 1701pp + Sequence Listing: English.
XX
XX
XX The invention relates to novel genes (ABA11004-ABA21354) and proteins
CC (ABA1678-ABA18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pctl_sequences.
XX
SQ Sequence 11641 BP; 3281 A; 2284 C; 2220 G; 3856 T; 0 other;
Alignment Scores:
Pred. No.: 1.16e+03 Length: 11641
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 22 Gaps: 0
US-09-727-892a-99 (1-58) x ABA14848 (1-11641)
OY 46 LysLysTyrAlaTyrIleIle 52
|||
Db 11518 AAAAAGTATCGTTATATTATA 11498
RESULT 42
AAK85387 standard; DNA: 11641 BP

XX AAK85387;
AC
XX
DT 07-NOV-2001 (first entry)
XX
DE Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:40199.
XX
KW Human; Immune; haematopoietic; Immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
OS Homo sapiens.
XX
PN W0200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WC-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.

PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.

PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 40199; 3071bp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK62170 to AAK91921. (I) have cytosstatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 11641 BP; 3856 A; 2220 C; 2284 G; 3281 T; 0 other;

Alignment Scores:
Pred. NO.: 1.16e+03 Length: 11641
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 22 Gaps: 0

US-09-727-892A-99 (1-58) x AAK65387 (1-11641)
QY 46 LysIstYrAlaTyrlleIle 52
Db 124 AAAAAGTATGCTTATATATA 144

RESULT 43
AAK65382
ID AAK65382 standard; DNA; 12904 BP.
XX
AC AAK65382;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/hematopoietic antigen genomic sequence SEQ ID NO:40194.
XX
KW Human; immune; haematopoietic; immune/hematopoietic antigen; cancer;
KW cytosstatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PM WO200157182-A2.
XX
PD 09-AUG-2001.

XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 26-JUL-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236359.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-023935.
PR 13-OCT-2000; 2000US-0239357.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249254.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251989.
PR 06-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251969.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX

Disclosure; SEQ ID NO 40194; 3071bp + Sequence Listing; English.
Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis -
AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I) amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patient's own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/hematopoietic-related diseases, especially cancers and cancer metastases of hematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/hematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAK82169 represent sequences used in the exemplification of the present invention.
Sequence 12904 BP; 4168 A; 2563 C; 2431 G; 3742 T; 0 other;
Alignment Scores:
Seq. No.: 1.27e+03 Length: 12904
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
B: 22 Gaps: 0
US-09-727-892A-99 (1-58) x AAK85382 (1-12904)
46 Lysylsyratyllelle 52
3880 AAAAGTATGCTTATTATA 3900
RESULT 44
ABL22282
ABL22282 standard; DNA; 13293 BP.
ABL22282;
26-MAR-2002 (first entry)
Drosophila melanogaster genomic polynucleotide SEQ ID NO 18319.
Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical; gene; ds.
Drosophila melanogaster.
WO200171042-A2.
27-SEP-2001.
23-MAR-2001; 2001WO-US09231.
23-MAR-2000; 2000US-191637P.
11-JUL-2000; 2000US-0614150.
(PEKE) PE CORP NY.
Venter JC, Adams M, Li PWD, Myers EM;
WPI; 2001-656860/75.
New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

XX Claim 1; SEQ ID NO 18319; 21bp + Sequence Listing; English.
PS
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB57737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 13293 BP; 4529 A; 2504 C; 2428 G; 3832 T; 0 other;
Alignment Scores:
Pred. No.: 1.31e+03 Length: 13293
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: Gaps: 0
US-09-727-892A-99 (1-58) x ABL22282 (1-13293)
QY 38 TYRTYRGLUTYrAsnLeuphe 44
DB 10501 TATTCAGAGTACACCTATTTC 10521
RESULT 45
AB133192/C
ID ABL33192 standard; DNA; 13814 BP.
XX
AC ABL33192;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 1165.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytosine methylation; antiasthmatic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PE 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
PS Claim 1; SEQ ID NO 1165; 32bp + Sequence Listing; German.
XX

CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 13814 BP; 3814 A; 288 C; 3168 G; 6543 T; 1 other;
Alignment Scores:
Pred. No.: 1.35e+03 Length: 13814
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: Gaps: 0
US-09-727-892A-99 (1-58) x ABL33192 (1-13814)
QY 41 TYrAsnLeuphenThrLysLys 47
DB 1903 TACAACCTCTTCACAAAAA 1883
Search completed: November 5, 2002, 04:22:40
Job time : 260 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 5, 2002, 04:19:49 : Search time 1716 Seconds
(without alignments)
456.191 Million cell updates/sec

Title: US-09-727-892A-99

Perfect score: 58

Sequence: 1 MERKKTIVLYXCDEIKGHF.....YEVNLPFRKKYATIEYIKEI 58

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 13736207 segs, 6748477542 residues

Word size: 1

Total number of hits satisfying chosen parameters: 27472246

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+lp2n.model -DEV-xlh
-O=cgn2.1/USPFO.spool/US09727882/ufunat_01112002_185930_4944/app-query.fasta.1.199
-DB=EST -QFWT=fastap -SUFFIX=NAolig.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR.SCORER=quality -THR.MIN=1 -ALIGN=45 -MODE=LOCAL -OUTFWT=ptio
-NORM=ext -HEARSQUE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09727882.ecgn.1.1-763-etunac_01112002_185930_4944 -NCPD=6 -ICPD=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WALT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

EST :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	13.8	266	9	AW719727 LJNEST9A1
2	8	13.8	282	10	NA5026 YY17A03.r1

c	3	13.8	321	9	BB403737	BB403737
c	4	13.8	337	9	BE232783	BE232783
c	5	13.8	355	9	AA097126	AA097126
c	6	13.8	466	9	AT142154	AT142154
c	7	13.8	511	10	BG798530	BG798530
c	8	13.8	510	10	BI862876	BI862876
c	9	13.8	539	10	BM131380	BM131380
c	10	13.8	560	10	BE927751	BE927751
c	11	13.8	616	9	AW783051	AW783051
c	12	13.8	630	12	A2525190	A2525190
c	13	13.8	702	12	BH520084	BH520084
c	14	13.8	963	10	BF030328	BF030328
c	15	12.1	157	10	BG549008	BG549008
c	16	12.1	161	10	T66327	T66327
c	17	12.1	164	9	AT186030	AT186030
c	18	12.1	167	10	F03862	F03862
c	19	12.1	169	12	BH122433	BH122433
c	20	12.1	171	9	AW770158	AW770158
c	21	12.1	173	10	BG957366	BG957366
c	22	12.1	180	9	AW628877	AW628877
c	23	12.1	180	10	C65613	C65613
c	24	12.1	180	10	C69034	C69034
c	25	12.1	181	10	D34512	D34512
c	26	12.1	182	9	AV352335	AV352335
c	27	12.1	182	10	T25932	T25932
c	28	12.1	184	10	R45661	R45661
c	29	12.1	198	10	D37587	D37587
c	30	12.1	201	10	N49179	N49179
c	31	12.1	208	9	AA613897	AA613897
c	32	12.1	209	12	CNS00WML	CNS00WML
c	33	12.1	211	10	F03788	F03788
c	34	12.1	222	12	A2080653	A2080653
c	35	12.1	228	9	AT681809	AT681809
c	36	12.1	232	10	F09939	F09939
c	37	12.1	234	10	F04344	F04344
c	38	12.1	250	9	AT766865	AT766865
c	39	12.1	258	10	F03130	F03130
c	40	12.1	256	9	AA353041	AA353041
c	41	12.1	264	9	AA732601	AA732601
c	42	12.1	267	10	BM424468	BM424468
c	43	12.1	270	9	BB287799	BB287799
c	44	12.1	272	12	AQ098659	AQ098659
c	45	12.1	275	9	AA975953	AA975953

ALIGNMENTS

RESULT 1
AW719727
LOCUS
DEFINITION
AW719727
ACCSSION
AW719727
VERSION
AW719727
KEYWORDS
SOURCE
ORGANISM
Lotus japonicus.
Lotus japonicus.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
Lotus.
REFERENCE
1 (bases 1 to 266)
Colebatch, G., Freund, S., Trevasakis, B and Udvardi, M.
Lotus japonicus root module ESTs: tools for functional genomics
unpublished (2000)
Contact: Udvardi, M.
Molecular Plant Nutrition
Max Planck Institute of Molecular Plant Physiology
Am Muehlenberg 1, 14476 Golm, Germany
Fax: 49 351 567 8250
Email: udvardi@mpimp-golm.mpg.de
Seq primer: 77
High quality sequence stop: 266.

Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers

1..321
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="C3003311"
/cell_type="ES cells"
/lab_host="SODR"

BASE COUNT

78 a 87 c 40 g 116 t

US-09-727-892a-99 (1-58) x BB403737 (1-321)

Alignment Scores:

Pred. No.: 34.5 Length: 321
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.79% Indels: 0
DB: 9 Gaps: 0

US-09-727-892a-99 (1-58) x BB403737 (1-321)

OY 9 Leuetyrcysaspjullelys 16

Db 175 CTATTGACTGTGATGAATAAA 152

RESULT 4

BE232783/c

LOCUS BE232783 337 bp mRNA linear EST 10-JUL-2000

DEFINITION 137999 MARC 1PTG Sus scrofa CDNA 5', mRNA sequence.

ACCESSION BE232783

VERSION BE232783.1 GI:9017501

KEYWORDS EST.

SOURCE pig.

ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

1 (bases 1 to 337)

Fahrenkrug,S.C., Feking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,

Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.M.

and Keele,J.W.

Design and use of two pooled tissue normalized CDNA libraries for

EST discovery in swine

Unpublished (2000)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt. trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 18

and -mismatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTTCCAGTCACGACG

Plate: 73 row: E column: 19

Seq primer: ATTAGGTGACACTATAG.

FEATURES

Location/Qualifiers

1..337
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_1lb="MARC 1PTG"
/tissue_type="pooled"
/lab_host="DHL08"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

BASE COUNT

97 a 48 c 77 g 115 t

US-09-727-892a-99 (1-58) x BE232783 (1-337)

Alignment Scores:

Pred. No.: 36.2 Length: 337
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.79% Indels: 0
DB: 9 Gaps: 0

US-09-727-892a-99 (1-58) x BE232783 (1-337)

OY 45 ThrlystlyrAlaTyrlle 52

Db 241 ACAAAGTATGCTTATATATA 218

RESULT 5

AA097126/c

LOCUS AA097126 355 bp mRNA linear EST 25-OCT-1996

DEFINITION 0275F Pyrococcus furiosus 1-ZAP II library, F Robb Pyrococcus

furiosus CDNA clone 0275, mRNA sequence.

ACCESSION AA097126

VERSION AA097126.1 GI:1642707

KEYWORDS EST.

SOURCE Pyrococcus furiosus

ORGANISM Pyrococcus furiosus

Archaea; Euryarchaeota; Thermococci; Thermococcales;

Thermococcaceae; Pyrococcus.

1 (bases 1 to 355)

Borges,K.M., Brummet,S.R., Bogert,A., Davis,M.C., Hujer,K.M., Domke

,S.T., Szasz,J., Ravel,J., Dirugliero,J., Fuller,C., Chase,J.W. and

Robb,F.T.

A Survey of the Genome of the Hyperthermophilic Archaeon,

Pyrococcus furiosus

Unpublished (1996)

Contact: Brummet SR; Robb F

Brummet, Sequencing R&D

Amerisham Life Science, Inc.

Tel: 2164649277; 4102348870

Fax: 2163600975; 4102348896

Email: cn288@cleveland.freenet.edu

Seq primer: T7.

FEATURES

Location/Qualifiers

1..355
/organism="Pyrococcus furiosus"
/strain="DSM 3638"
/db_xref="taxon:2261"
/clone="0275"
/clone_1lb="Pyrococcus furiosus 1-ZAP II library, F Robb"
/lab_host="E. coli"
/note="Vector: 1-ZAP II; Site_1: Eco RI; Site_2: Eco RI;
Genomic DNA was purified from P. furiosus cells and
partially digested with Dra I, Eco RV, Hinc II, Pvu II,
size selected, ligated to Eco RI linkers then cloned into
the Eco RI site of 1-ZAP II, plasmid excision vector.
Excision was performed in batch and individual clones
retrieved by plating."

BASE COUNT

104 a 80 c 120 t 11 others

ORIGIN

Alignment Scores:

Pred. No.:	38.2	Length:	355
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	13.79%	Indels:	0
DB:	9	Gaps:	0

US-09-727-892a-99 (1-58) x AA097126 (1-355)

OY 23 IleserMethegluaspleuty 30
 DB 202 ATTTCATGTTCAGATTTGTAT 179

RESULT 6
 A1142154 466 bp mRNA linear EST 02-OCT-2000
 LOCUS A1142154/c
 DEFINITION EST042 Manduca sexta male antennae Uni-ZAP XR library Manduca sexta
 ACCESSION A1142154
 VERSION A1142154.1 GI:3658513
 KEYWORDS EST.
 SOURCE tobacco hornworm.
 ORGANISM Manduca sexta
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia
 ; Spingiodae; Spingidae; Sphinginae; Manduca.
 1 (bases 1 to 466)
 Robertson, H.M., Martos, R., Sears, C.R., Todres, E.Z., Walden, K.K. and
 Nardl, J.B.
 Diversity of odourant binding proteins revealed by an expressed
 sequence tag project on male Manduca sexta moth antennae
 Insect Mol. Biol. 8, 501-518 (1999)
 20099029
 COMMENT Contact: Robertson HM
 Department of Entomology
 University of Illinois at Urbana-Champaign
 505 S. Goodwin, Urbana, IL 61801, USA
 Tel: 217 333-0489
 Fax: 217 244 3499
 Email: hughro@uiuc.edu
 Insert Length: 1300 Std Error: 0.00
 Seq primer: SK
 High quality sequence stop: 400.
 Location/Qualifiers
 1..466
 /organism="Manduca sexta"
 /db_xref="taxon:7130"
 /clone_lib="Manduca sexta male antennae Uni-ZAP XR
 library"
 /sex="male"
 /dev_stage="newly eclosed adults and pharate adults"
 /lab_host="X11 Blue MFR and SOLR"
 /note="Organ: antennae; Vector: Uni-ZAP XR. Site: 1: EcoRI;
 Site: 2: XhoI; The library was prepared by StrataGene using
 oligo-T priming and unidirectional cloning with an
 adaptor at the 5' end (GGCAGCAG) following the EcoRI
 site. The mRNA was prepared from antennae of late pupal
 and newly eclosed male moths. Clones were subcloned in
 vivo in mass into pBluescript maintained in SOLR cells
 for DNA sequencing."
 BASE COUNT 133 a 112 c 75 g 146 t
 ORIGIN

Alignment Scores:

Pred. No.:	50.2	Length:	466
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	13.79%	Indels:	0
DB:	9	Gaps:	0

US-09-727-892a-99 (1-58) x A1142154 (1-466)

OY 36 TySerTyTyGrUtyrAsnLeu 43
 DB 143 TACAGCTATTCATATATATTTA 120

RESULT 7
 BG798530 510 bp mRNA linear EST 05-JUL-2001
 LOCUS BG798530
 DEFINITION BG798530
 ACCESSION BG798530.1 GI:14162862
 VERSION BG798530.1 GI:14162862
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 510)
 Melton, D., Brown, J., Kenly, G., Permut, A., Lee, C., Kaestner, K.,
 Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
 Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Bistacchi, A.,
 Schmitt, A., Theising, B., Ritzer, E., Ronko, I., Bennett, J., Cardenas
 M., Gibbons, M., McCann, R., Cole, R., Tsagaris, R., Williams, T.,
 Jackson, Y. and Bowers, Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)
 Other ESTs: ic10f10.x1
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biochem.harvard.edu
 Pancreas was obtained from Gerard Gradwohl (PNAS 97 p1607-1611,
 2000) Library was constructed by Catherine Lee DNA sequencing by:
 Washington University Genome Sequencing Center for information on
 obtaining a clone please contact: Dr. Marie Searce
 (mscarce@mail.med.upenn.edu)
 High quality sequence stop: 436.
 Location/Qualifiers
 1..510
 /organism="Mus musculus"
 /strain="129/Sv x CD1"
 /db_xref="taxon:10090"
 /clone_lib="Kaestner ngn3 - -"
 /dev_stage="p.c. 14.5"
 /lab_host="E. coli-DH12S (GIBCO)"
 /note="Organ: pancreas; Vector: pSPORT2 (GIBCO); Site: 1:
 Not I; Site: 2: Sal I; The library was prepared by
 Catherine S. Lee and has not been published. The pancreas
 was obtained from Gerard Gradwohl (PNAS 97 p1607-1611,
 2000). The cDNA's were prepared with an oligo containing a
 NotI site, and SalI linkers were added to the ends. The
 inserts were cut with NotI before being cloned into the
 NotI-SalI sites in the vectors. This is one of two
 libraries, ngn3 wt and ngn3 -/-. The ngn3 -/- library is
 in pSPORT2, T7 promoter is 3'."
 BASE COUNT 166 a 86 c 83 g 175 t
 ORIGIN

Alignment Scores:

Pred. No.:	55	Length:	510
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	13.79%	Indels:	0
DB:	10	Gaps:	0

US-09-727-892a-99 (1-58) x BG798530 (1-510)

OY 40 GlutyrAsnLeuPheThrLysLys 47

Db 131 GAATATATTTATTATCCCAAAA 154

RESULT 8

BI862876/c 511 bp mRNA linear EST 10-OCT-2001
 LOCUS rml3e03.y1 Meloidogyne arenaria egg PAMP1 v1 Chiapelli McCarter
 DEFINITION Meloidogyne arenaria cDNA 5' similar to TR:002353 002353

3-KETO-ACYL-COA THIOLASE. [1] ; mRNA sequence.

ACCESSION BI862876 GI:16004666

VERSION EST.

KEYWORDS Meloidogyne arenaria.

SOURCE Meloidogyne arenaria

ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;

REFERENCE Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.

AUTHORS

1 (bases 1 to 511)
 McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J., Wylie, T.,
 Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y.,
 Gibbons, M., Rittler, E., Bennett, J., Franklin, C., Tsagareishvili, R.,
 Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe,
 M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurr, R., Kohn, S.,
 Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and
 Wilson, R.

Unpublished (1999)
 The Washington Univ. Nematode EST Project, 1999

CONTACT: McCarter JP

The Washington Univ. Nematode EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

The library was constructed by Brandi Chiapelli and Dr. James
 McCarter (bchiapelle@watson.wustl.edu & jmcarter@watson.wustl.edu) at
 Washington University, St. Louis. DNA sequencing by: Washington
 University Genome Sequencing Center St. Louis. Nematodes were
 provided by Dr. David Bird and Daniel Snyder of North Carolina
 State University.
 Seq primer: -40RP from GIBCO
 High quality sequence stop: 415.

FEATURES

source

1..511
 Location/Qualifiers

1..511

/organism="Meloidogyne arenaria"

/db_xref="taxon:6304"

/clone_id="Meloidogyne arenaria egg PAMP1 v1 Chiapelli

McCarter"

/dev_stage="egg"

/lab_host="DH10B"

/note="Vector: PAMP1 (Gibco); Site_1: NotI; Site_2: SalI;

The library was constructed by Brandi Chiapelli and Dr.

James McCarter at Washington University, St. Louis. The

cDNA was made by using Dynabead oligo-dT priming (Dyna).

PCR based library using a modified protocol from the

SMART PCR cDNA Synthesis Kit from Clontech. Directionally

cloned into the UDG sites of PAMP1. Nematodes were

provided by Dr. David Bird and Daniel Snyder of North

Carolina State University."

BASE COUNT

174 a 63 c 95 g 179 t

ORIGIN

Alignment Scores:

Pred. No.: 55.1 Length: 511
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 13.79% Indels: 0
 DB: 10 Gaps: 0

US-09-727-892a-99 (1-58) x BI862876 (1-511)

QY 41 TyranLeuphethrlystlyr 48

|||||

Db 466 TATATTTGTTTACAAAGAAATAT 443

RESULT 9

BM131380 539 bp mRNA linear EST 27-NOV-2001
 LOCUS TgESTzYa93g06.y1 TgVEG Partially sporulated oocyst cDNA Toxoplasma
 DEFINITION gondii cDNA clone TgESTzYa93g06.y1 5', mRNA sequence.

ACCESSION BM131380 GI:17125932

VERSION EST.

KEYWORDS Toxoplasma gondii.

SOURCE Toxoplasma gondii

ORGANISM Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;

Sarcocystidae; Toxoplasma.

REFERENCE 1 (bases 1 to 539)

Tang, K., Cole, R., Fogarty, S., Sibley, L.D., Ajioke, J.A., White, M.,
 Clifton, S., Pape, D., Martin, J., Wylie, T., Dante, M., Marra, M.,
 Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Rittler,
 E., Bennett, J., Franklin, C., Tsagareishvili, R., Ronko, I., Kennedy,
 S., Maguire, L., Waterston, R. and Wilson, R.

Toxoplasma EST Project

Unpublished (2001)

Contact: Clifton, S.

Toxoplasma EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: toxo@watson.wustl.edu

Contact David Sibley (dsibley@watson.wustl.edu) for further

information relating to organism, libraries, or clone availability.

Seq primer: -40RP from GIBCO

High quality sequence stop: 439.

Location/Qualifiers

1..539

/organism="Toxoplasma gondii"

/strain="VEG (Type III)"

/db_xref="taxon:5811"

/clone="TgESTzYa93g06.y1"

/clone_id="TgVEG Partially sporulated oocyst cDNA"

/dev_stage="Partially sporulated oocysts"

/note="Vector: Modified pBluescript (pBSSK+); Site_1:

BamHI; Site_2: EcoRI; PolyA mRNA from partially-sporulated

oocysts was converted to cDNA using the

template-switching PCR method (SMART cDNA, Clontech Inc.,)

and sized selected on SizeSeph 400 columns (Amersham

pharmacia Biotech Inc.). First strand was reverse

transcribed using the CDS III-oligo-dT primer and a 5'

template switch primer (Smart IV primer). The product of

the first strand synthesis was PCR amplified using the

same primer set and the fragments were digested with SfiI.

The fragments were size selected, ligated into a modified

pBluescript vector containing directional SfiI sites, and

electroporated into DH10B or DH12S cells. Vector: SfiI

sites were added to the multiple cloning region of

pBluescript SK+ between the BamHI/EcoRI sites. The

modified polylinker has the following sequence:

5'GAAATCGGCATTTACGGCC(G)n--Insert--

GGCCGCTCGGCCCGCCGATGTC3'where n=3-4 G nucleotides. Library

Source: Michael White, Maria E. Jerome, Emily A. Johnson,

Jay A. Radke, Montana State University. Clone

Availability: David Sibley, Washington University"

BASE COUNT

130 a 117 c 137 g 155 t

ORIGIN

Alignment Scores:

Pred. No.: 58.2 Length: 539
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 13.79% Indels: 0
 DB: 10 Gaps: 0

US-09-727-892a-99 (1-58) x BM131380 (1-539)
 OY 3 ArglystylyrPhrVallieu10
 ||||||||||||||||||
 DB 474 CGGAGATATAGACCGTCTTTTG 451

RESULT 10
 BE927751 560 bp mRNA linear EST 02-OCT-2000
 LOCUS MR3-CT0462-250800-003-g05 CT0462 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BE927751
 ACCESSION BE927751
 VERSION BE927751.1 GI:10453827
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 560)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 CONTACT: Simpson A.J.G.
 LABORATORY OF Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-MR3-CT0462-250
 800-003-g05&ts=2000-08-25&tl=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 560.

FEATURES

source
 1. 560
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="CT0462"
 /dev_stage="Adult"
 /note="Organ: colon; Vector: puc18; Site:1: Smal; Site:2:
 Sma1; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No. 196
 /716 - Ludwig Institute for Cancer Research) profiles
 into the puc 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."
 BASE COUNT 189 a 106 c 88 g 176 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 60.5 Length: 560
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 13.79% Indels: 0
 DB: 10 Gaps: 0

US-09-727-892a-99 (1-58) x BE927751 (1-560)

OY 45 ThrlslystylyrAlarYlle1e 52
 ||||||||||||||||||
 DB 6 ACAAAAAGTATGCTTACATTATA 29
 RESULT 11

AM783051/c 616 bp mRNA linear EST 10-MAY-2001
 LOCUS ra22q08.y1 Bird-Rao Meloidogyne incognita J2 Meloidogyne incognita
 DEFINITION CDNA 5' similar to WP.Y57A10C.6 CEI8418 3-KETO-ACYL-COA THIOLASE ;
 mRNA sequence.
 ACCESSION AM783051
 VERSION AM783051.1 GI:7797657
 KEYWORDS EST.
 SOURCE southern root-knot nematode.
 ORGANISM Meloidogyne incognita
 Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchida;

REFERENCE 1 (bases 1 to 616)
 McCarter,J., Clifton,S., Chiappelli,B., Pape,D., Martin,J., Wylie,T.,
 Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.,
 Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagarisylli,R.,
 Ronko,T., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Streptoe
 M., Allen,M., Person,B., Swaller,T., Harvey,N., Schutk,R., Kohn,S.,
 Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
 Wilson,R.
 The Washington Univ. Nematode EST Project, 1999
 Unpublished (1999)
 CONTACT: McCarter JP
 The Washington Univ. Nematode EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 The library was constructed by Uma Rao and David Bird
 (david_bird@ncsu.edu) at North Carolina State University. DNA
 Sequencing by: Washington University Genome Sequencing Center St.
 Louis.

Seq primer: T3 ET from Amersham
 High quality sequence stop: 430.
 Location/Qualifiers
 1. 616
 /organism="Meloidogyne incognita"
 /db_xref="taxon:6306"
 /clone_id="Bird-Rao Meloidogyne incognita J2"
 /dev_stage="enriched for 2nd stage juveniles"
 /lab_host="XtRL"

/note="Vector: ZAP express - pBKCMV (Stratagene); Site:1:
 EcoRI; Site:2: XhoI; Oligo (dT) primed library. cDNA was
 constructed and cloned unidirectionally into the vector
 within the 5' EcoRI and 3' XhoI sites. This library was
 constructed by Dr. Uma Rao and Dr. David Bird at North
 Carolina state University."
 BASE COUNT 216 a 94 c 123 g 182 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 66.6 Length: 616
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 13.79% Indels: 0
 DB: 9 Gaps: 0

US-09-727-892a-99 (1-58) x AM783051 (1-616)

OY 41 Tyranleupherthlystylyr 48
 ||||||||||||||||||
 DB 394 TATAATTGTTTACAAAGAAATAT 371
 RESULT 12
 A2525190/c 630 bp DNA linear GSS 07-MAY-2001
 LOCUS A2525190
 DEFINITION 241PB12 Pb MBN #21 Plasmodium berghel genomic 3', DNA sequence.
 ACCESSION A2525190
 VERSION A2525190.1 GI:13965795
 KEYWORDS GSS.
 SOURCE Plasmodium berghel.

ORGANISM Plasmodium berghei
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE 1 (bases 1 to 630)
AUTHORS Carlton,J.M.-R. and Dame,J.B.
TITLE The Plasmodium vivax and P. berghei gene sequence tag projects
JOURNAL Parasitol. Today 16 (10), 409 (2000)
COMMENT

Dept. of Pathobiology, College of Veterinary Medicine
University of Florida
2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
Tel: 352 392 4700
Fax: 352 392 9704
Email: damej@mail.vetmed.ufl.edu
Seq primer: M13(-20) forward
Class: shotgun.

FEATURES
source
1. 630
/organism="Plasmodium berghei"
/strain="ANKA clone 15cyl (clone of the ANKA 8417 clone)"
/db_xref="taxon:5821"
/clone_1lb="Pb MEN #21"
/dev_stage="asexual blood forms"
/lab_host="Mus musculus"
/note="Vector: pBluescript SK(+) vector DNA, phagemid excised from lambda ZAP; Site.1: EcoRV; Site.2: EcoRV; Genomic DNA was prepared from asynchronous blood stage forms of the cloned ANKA isolate of P. berghei grown in laboratory Swiss white mice. The DNA was purified from ultracentrifugation and precipitated. Purified DNA was digested with mung bean nuclease in the presence of 36-38% formamide at 50 C, as described (Vernick, K.D., Imberski, R.B., and McCutchan, T.F. 1988. Nucleic Acids Research 16:6883-6896). The ends of the digestion fragments were polished using T4 DNA polymerase, and the fragments size selected in the range 500-2000 bp. These were ligated into the EcoRV-cleaved and dephosphorylated pBluescript SK(+) vector. Recombinant plasmids were used to transform E. coli XL10-Gold host cells."

BASE COUNT 293 a 48 c 77 g 212 t
ORIGIN

Alignment Scores:
Pred. No.: 68.1 Length: 630
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.79% Indels: 0
Gaps: 0

US-09-727-892a-99 (1-58) x AZ525190 (1-630)

QY 44 PheThrLysLysTyrAlaTyrIle 51
Db 64 TTCACAAAATAATGATCATACATA 41

RESULT 13
BH520084
LOCUS BH520084 702 bp DNA linear GSS 13-DEC-2001
DEFINITION BOGR41R BOGR Brassica oleracea genomic clone BOGR41, DNA sequence.
ACCESSION BH520084
VERSION BH520084.1 GI:17728169
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 702)
AUTHORS Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)

COMMENT Other-GSSs: BOGR41R
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES
source
1. 702
/organism="Brassica oleracea"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone_1lb="BOGR41"
/clone_1lb="BOGR"
/note="Vector: pHOS1; Site.1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"

BASE COUNT 237 a 150 c 114 g 201 t
ORIGIN

Alignment Scores:
Pred. No.: 76 Length: 702
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.79% Indels: 0
Gaps: 0

US-09-727-892a-99 (1-58) x BH520084 (1-702)

QY 41 TyrAsnLeuPheThrLysLysTyr 48
Db 218 TATATCTATTATACCAAGAACTAC 241

RESULT 14
BF030328/c
LOCUS BF030328/c 963 bp mRNA linear EST 10-OCT-2000
DEFINITION 601558374F1 NIH_MGC_58 Homo sapiens CDNA clone IMAGE:3828043 5', mRNA sequence.
ACCESSION BF030328
VERSION BF030328.1 GI:10738040
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 963)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LICM498 row: C column: 20
High quality sequence stop: 305.

FEATURES
source
1. 963
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="IMAGE:3828043"
/clone_1lb="NIH_MGC_58"
/tissue_type="hypertrophoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site.1: SfiI (ggccgctcgcc); Site.2: SfiI (ggccattatggcc);

US-09-727-892a-99 (1-58) x T66327 (1-161)

OY 46 LyslystYrAlaTyrllelle 52
 |||||||||||||||||||
 Db 19 AAAAAGTATGCTTATATATA 39

RESULT 17
 AUI86030/c 164 bp mRNA linear EST 05-JUL-2001
 LOCUS AUI86030 Homo sapiens T-cell library (Sugita Y) Homo sapiens cDNA
 DEFINITION clone B02534-021, mRNA sequence.
 ACCESSION AUI86030
 VERSION AUI86030.1 GI:14623943
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 164)
 AUTHORS Sugita,Y., Oshida,T. and Oya,Y.
 TITLE Human cDNA sequencing
 JOURNAL Unpublished (2001)
 COMMENT Contact: Yuji Sugita
 Genex Research, Inc.
 907 Nogawa, Miyamae-ku, Kawasaki, Kanagawa 216-0001, Japan
 Tel: 81-44-797-2281
 Fax: 81-44-797-2622
 Email: syuji@genex.co.jp, URL: http://www.genex.co.jp.

FEATURES
 source
 1..164
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="B02534-021"
 /clone_lib="Homo sapiens T-cell library (Sugita Y)"
 /cell_type="T-cell"

BASE COUNT 50 a 31 c 20 g 63 t

ORIGIN

Alignment Scores:
 Pred. No.: 211 Length: 164
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 Gaps: 0

US-09-727-892a-99 (1-58) x AUI86030 (1-164)

OY 18 HisPheProHisGlnIleSer 24
 |||||||||||||||||||
 Db 73 CATTTCACATCATGATTAAGC 53

RESULT 18
 F03862 167 bp mRNA linear EST 19-FEB-1995
 LOCUS F03862 HSC2CG042 normalized infant brain cDNA Homo sapiens cDNA clone
 DEFINITION c-2cB04 3', mRNA sequence.
 ACCESSION F03862
 VERSION F03862.1 GI:670486
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 167)
 AUTHORS Aufiray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes
 M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,
 Mitchell,H., Mariage-Samson,R., Pletu,G., Pouliot,Y.,
 Sebastiant-Kabaktchis,C. and Tessier,A.
 TITLE IMAGE: molecular integration of the analysis of the human genome
 and its expression
 JOURNAL C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
 MEDLINE 95277534

COMMENT
 Contact: Genethon
 Genexpress-Genethon
 Genethon Centre de recherche sur le Genome Humain
 1,rue de l'Internationale, Bp60 91002 EVRY Cedex, FRANCE
 Tel: 33169472800
 Fax: 33160778698
 Email: genexpress@genethon.fr
 Single read removed at sequence 5' end
 Genexpress_library_id: C; Genexpress_sequence_id: alc-2cB04
 Seq primer: (-21)M13-universal.
 Location/Qualifiers
 1..167
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="c-2cB04"
 /clone_lib="normalized infant brain cDNA"
 /sex="Female"
 /tissue_type="total brain"
 /dev_stage="3 months old"
 /note="Organ: brain; Vector: lafmid BA; Site_1: HindIII;
 Site_2: NotI; sex=Female; dev_stage=3 months old;
 Isolate=muscular atrophy patient; tissue_type=total brain
 ; total mRNA was oligo-(dt) primed and directionally
 cloned 5' -> 3' into the HindIII -> NotI sites of the
 lafmid BA vector. Clone library from B.Souares, Psychiatry
 Dept. Columbia University, USA
 Bento Soares, P.N.A.S in press"

BASE COUNT 62 a 24 c 19 g 54 t 8 others

ORIGIN

Alignment Scores:
 Pred. No.: 215 Length: 167
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 Gaps: 0

US-09-727-892a-99 (1-58) x F03862 (1-167)

OY 46 LyslystYrAlaTyrllelle 52
 |||||||||||||||||||
 Db 99 AAAAAGTATGCTTATATATA 119

RESULT 19
 BH122433 169 bp DNA linear GSS 19-JUL-2001
 LOCUS BH122433 RPCI-24-30017.TV RPCI-24 Mus musculus genomic clone RPCI-24-30017,
 DEFINITION DNA sequence.
 ACCESSION BH122433
 VERSION BH122433.1 GI:14965945
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 169)
 AUTHORS Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akintet,B., Levins,M.,
 Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregorgis,E.,
 Russell,D., de Jong,P. and Fraser,C.M.
 Mouse BAC End Sequences from Library RPCI-24
 Unpublished (1999)
 CONTACT: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhaoe@tigr.org
 Clones are derived from the mouse BAC library RPCI-24. For BAC
 library availability, please contact Pieter de Jong
 (pdejong@mail.cho.org). Clones may be purchased from BACPAC
 Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end

page: http://www.tigr.org/tdb/bac-ends/mouse/bac-end_intro.html
Plate: 300 row: 1 column: 7

Seq primer: T7
Class: BAC ends.

FEATURES

source

Location/Qualifiers
1. 169

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-30017"
/clone_1lb="RPCI-24"
/sex="Male"

/cell_type="Spleen/Brain"
/note="Vector: pTRABAC1; Site_1: BamHI; Site_2: BamHI;
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTRABAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."

BASE COUNT 52 a 31 c 27 g 59 t
ORIGIN

Alignment Scores:

Pred. No.: 217 Length: 169
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 12 Gaps: 0

US-09-727-892a-99 (1-58) x BH122433 (1-169)

OY 31 Aspalalyvaltyrser 37

Db 87 GATCAAAAGTACTTACACT 107

RESULT 20

AM770158 171 bp mRNA linear EST 04-MAY-2000

LOCUS AM770158 171 bp mRNA linear EST 04-MAY-2000
DEFINITION h559a06.x1 NCI-CGAP-Lym12 Homo sapiens cDNA IMAGE:3000946 3',
mRNA sequence.

ACCESSION AM770158
VERSION AM770158.1 GI:7702197

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov

Life Technologies catalog #: 11547-015
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
infoimage.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 151.
Location/Qualifiers

FEATURES
source

1. 171

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3000946"
/clone_1lb="NCI-CGAP-Lym12"
/tissue_type="Lymphoma, follicular mixed small and large
cell"

/lab_host="DH10B"
/note="Organ: lymph node; Vector: pCMV-SPORT6; Site_1:
SalI; Site_2: NotI; Cloned unidirectionally. Primer:
Oligo dT. Average insert size 1.25 kb. Life Technologies

BASE COUNT 62 a 23 c 17 g 69 t
ORIGIN

Alignment Scores:

Pred. No.: 220 Length: 171
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 9 Gaps: 0

US-09-727-892a-99 (1-58) x AM770158 (1-171)

OY 46 Lyslystyraltyrlelle 52

Db 117 AAAAAGTATGCTATATATA 137

RESULT 21

BG957366/c 173 bp mRNA linear EST 12-JUN-2001

LOCUS BG957366 173 bp mRNA linear EST 12-JUN-2001
DEFINITION IL3-CT0674-010301-490-B07 CT0674 Homo sapiens cDNA, mRNA sequence.

ACCESSION BG957366
VERSION BG957366.1 GI:14375537

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 173)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagal, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663

JOURNAL

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3-CT0674-
010301-490-B07&ts=2001-03-01&ta=1)

Seq primer: puc 18 forward
High quality sequence start: 22
High quality sequence stop: 173.
Location/Qualifiers

FEATURES

source

1. 173

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="CT0674"
/dev_stage="Adult"

/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 42 a 35 c 30 g 66 t
ORIGIN

Alignment Scores: 223 Length: 173
Pred. No.: 223

Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 DB: 10 Gaps: 0

US-09-727-892a-99 (1-58) x BG957366 (1-173)

OY 34 ValValTYrSerTYrTYrGLU 40
 DB 95 GTGGTATTTCATCTATGAA 75

RESULT 22

LOCUS AM628877 180 bp mRNA linear EST 31-MAR-2000
 DEFINITION h144d03.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
 IMAGE:2975141 3, mRNA sequence.

ACCESSION AM628877
 VERSION AM628877.1 GI:7375667

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS 1 (bases 1 to 180)
 NCBI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: c9apds-r@mail.nih.gov

This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (infoimage.lnl.gov) for further information.
 Seq primer: -40UP from Gibco.
 High quality sequence stop: 172.

FEATURES
 source Location/Qualifiers

1..180
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2975141"
 /clone_id="Soares_NFL_T_GBC_S1"
 /lab_host="DH10B"
 /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
 a modified polylinker; Site 1: Not I; Site 2: Eco RI;
 Equal amounts of plasmid DNA from three normalized
 libraries (fetal lung NBHL19W, testis NHT, and B-cell
 NCI-CCAP_GCB1) were mixed, and ss circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The driver
 was PCR-amplified cDNAs from pools of 5,000 clones made
 from the same 3 libraries. The pools consisted of
 I.M.A.G.E. clones 297480-302087, 682632-687239,
 726408-726711, and 729096-731399. Subtraction by Bento
 Soares and M. Fatima Bonaldo."

BASE COUNT 75 a 26 c 20 g 59 t

ORIGIN

Alignment Scores: 232 Length: 180
 Pred. No.: 7.00 Matches: 7
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 12.07% Indels: 0
 Query Match: 12.07% Gaps: 0

US-09-727-892a-99 (1-58) x AM628877 (1-180)

OY 46 LysLysTYrAlaTYrIleIle 52
 DB 101 AAAAAGTATGCTTATATATA 121

RESULT 23
 C65613

LOCUS C65613 180 bp mRNA linear EST 22-SEP-1997
 DEFINITION C65613 Yuiji Kohara unpublished cDNA Caenorhabditis elegans cDNA
 clone yk429d8 5', mRNA sequence.

ACCESSION C65613
 VERSION C65613.1 GI:2424318

KEYWORDS EST.

SOURCE Caenorhabditis elegans.

ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
 ; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano
 1 (bases 1 to 180)

AUTHORS 'M., Miyata,A. and Nishigaki,A.
 Expression map of the C.elegans genome

TITLE Unpublished (1996)

JOURNAL Contact: Yuiji Kohara
 COMMENT Genome Biology Lab.
 National Institute of Genetics
 Yata 1111, Mishima, Shizuoka 411, Japan
 Tel: 81-559-81-6854
 Fax: 81-559-81-6855
 Email: ykohara@lab.nig.ac.jp.

FEATURES
 source Location/Qualifiers

1..180
 /organism="Caenorhabditis elegans"
 /strain="CB1489 him-8(e1489)"
 /db_xref="taxon:6239"
 /clone="yk429d8"
 /clone_id="Yuiji Kohara unpublished cDNA"
 /sex="hermaphrodite, male"
 /tissue_type="whole animal"
 /dev_stage="varied"

BASE COUNT 59 a 36 c 37 g 44 t 4 others

ORIGIN

Alignment Scores: 232 Length: 180
 Pred. No.: 7.00 Matches: 7
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 12.07% Indels: 0
 Query Match: 12.07% Gaps: 0

US-09-727-892a-99 (1-58) x C65613 (1-180)

OY 18 HisPheProHisGlnIleSer 24

DB 104 CACTTCCTCACCAGATTTC 124

RESULT 24

LOCUS C69034 180 bp mRNA linear EST 23-SEP-1997
 DEFINITION C69034 Yuiji Kohara unpublished cDNA Caenorhabditis elegans cDNA
 clone yk299f12 5', mRNA sequence.

ACCESSION C69034
 VERSION C69034.1 GI:2430390

KEYWORDS EST.

SOURCE Caenorhabditis elegans.

ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
 ; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano
 1 (bases 1 to 180)

AUTHORS 'M., Miyata,A. and Nishigaki,A.
 Expression map of the C.elegans genome

TITLE Unpublished (1996)

JOURNAL Contact: Yuiji Kohara
 COMMENT Genome Biology Lab.
 National Institute of Genetics
 Yata 1111, Mishima, Shizuoka 411, Japan
 Tel: 81-559-81-6854
 Fax: 81-559-81-6855
 Email: ykohara@lab.nig.ac.jp.

FEATURES Location/Qualifiers
 source 1.180
 /organism="Caenorhabditis elegans"
 /strain="CB1489 him-8(e1489)"
 /db_xref="taxon:6239"
 /clone="YK299f12"
 /clone_1lb="Yui Kohara unpublished cDNA"
 /sex="hermaphrodite, male"
 /tissue_type="whole animal"
 /dev_stage="varied"

BASE COUNT 57 a 38 c 36 g 44 t 5 others
 ORIGIN

Alignment Scores:
 Pred. No.: 232 Length: 180
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 DB: 10 Gaps: 0

US-09-727-892a-99 (1-58) x C69034 (1-180)

QY 18 Hispherohsgn1leser 24
 Db 112 CACTCCCTCCACGAGATTCT 132

RESULT 25
 D34512/c 181 bp mRNA linear EST 05-AUG-1994
 LOCUS CELK004C1R Yui Kohara unpublished cDNA Caenorhabditis elegans CDNA
 DEFINITION clone yk4c1 3', mRNA sequence.
 ACCESSION D34512
 VERSION D34512.1 GI:522544
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans.
 ORGANISM Caenorhabditis elegans.
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae.
 1 (bases 1 to 181)
 Kohara,Y., Mitsuki,H., Nishigaki,A., Motomashi,T., Sugimoto,A. and Tabara,H.
 Toward an expression map of the C.elegans genome
 Unpublished (1994)
 CONTACT: Yui Kohara
 Genome Biology Lab.
 National Institute of Genetics
 Yata 1111, Mishima, Shizuoka 411, Japan
 Tel: 81-559-81-6854
 Fax: 81-559-81-6855
 Email: ykohara@lab.nig.ac.jp.
 Location/Qualifiers
 1.181
 /organism="Caenorhabditis elegans"
 /strain="CB1489 him-8(e1489)"
 /db_xref="taxon:6239"
 /clone="YK4C1"
 /clone_1lb="Yui Kohara unpublished cDNA"
 /sex="hermaphrodite, male"
 /tissue_type="whole animal"
 /dev_stage="varied"

BASE COUNT 65 a 21 c 38 g 56 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 233 Length: 181
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 DB: 10 Gaps: 0

US-09-727-892a-99 (1-58) x D34512 (1-181)

FEATURES Location/Qualifiers
 source 1.182
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="7030406K18"
 /clone_1lb="RIKEN full-length enriched, 11 days embryo gonad"
 /sex="mixed"
 /tissue_type="gonad"
 /dev_stage="11 days embryo"
 /lab_host="DH10B"
 /note="Site 1: Sali; Site 2: BamHI. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'

RESULT 26
 AV352335/c 182 bp mRNA linear EST 12-NOV-1999
 LOCUS AV352335 RIKEN full-length enriched, 11 days embryo gonad Mus
 DEFINITION musculus cDNA clone 7030406K18 3', mRNA sequence.
 ACCESSION AV352335
 VERSION AV352335.1 GI:6393392
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 182)
 Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Takahashi,F., Tateo,M., Tomioka,N., Tsunoda,Y., Watanishi,A., Watanabe,S., Yamamura,T., Yasunishi,A., Yokota,T., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Konno,H., et al. 1999)
 Unpublished (1999)
 CONTACT: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-resgsc.riken.go.jp/
 URL: http://genome.gsc.riken.go.jp/
 Sasaki,N., Izawa,M., Watanishi,M., Ozawa,K., Tanaka,T., Yoneda,Y., Matsura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
 'Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
 Itoh,M., Kitsuai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
 Carninci,P. and Hayashizaki,Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
 Please visit our web site (http://genome.riken.go.jp) for further details.
 Location/Qualifiers
 1.182
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="7030406K18"
 /clone_1lb="RIKEN full-length enriched, 11 days embryo gonad"
 /sex="mixed"
 /tissue_type="gonad"
 /dev_stage="11 days embryo"
 /lab_host="DH10B"
 /note="Site 1: Sali; Site 2: BamHI. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'

US-09-727-892a-99 (1-58) x R45661 (1-184)

Oy 46 LysLysTYRAlATYrIleIle 52
|||||

Db 111 AAAAAGTATGCTTATTATATA 131

RESULT 29

D37587

LOCUS 198 bp mRNA linear EST 08-AUG-1994

DEFINITION CELK004C1F Yuj1 Kohara unpublished CDNA Caenorhabditis elegans CDNA

ACCESSION D37587

VERSION D37587.1 GI:525971

KEYWORDS EST.

SOURCE Caenorhabditis elegans.

ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea

REFERENCE 1 (bases 1 to 198)

AUTHORS Kohara,Y., Mitsuki,H., Nishigaki,A., Motohashi,T., Sugimoto,A. and

TABARA,H.

TITLE Toward an expression map of the C.elegans genome

JOURNAL Unpublished (1994)

COMMENT Contact: Yuj1 Kohara

Genome Biology Lab.

National Institute of Genetics

Yata 1111, Mishima, Shizuoka 411, Japan

Tel: 81-559-81-6854

Fax: 81-559-81-6855

Email: ykohara@lab.nig.ac.jp.

FEATURES Location/Qualifiers

1..198

/organism="Caenorhabditis elegans"

/strain="CB1489 him-8(e1489)"

/db_xref="taxon:6239"

/clone="YK4C1"

/clone_id="Yuj1 Kohara unpublished CDNA"

/sex="hermaphrodite, male"

/tissue_type="whole animal"

/dev_stage="varied"

BASE COUNT 66 a 40 c 24 g 68 t

ORIGIN

Alignment Scores:

Pred. No.: 255 Length: 198

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 12.07% Indels: 0

DB: 10 Gaps: 0

US-09-727-892a-99 (1-58) x D37587 (1-198)

Oy 7 ThYValleuLeuTYrCYsAsp 13

Db 117 ACgGTCTCTGCTATATGCGAT 137

RESULT 30

N49179

LOCUS 201 bp mRNA linear EST 14-FEB-1996

DEFINITION yy84606.s1 Soares,multiple_sclerosis_2NBHMSF Homo sapiens CDNA

ACCESSION N49179

VERSION N49179.1 GI:1190345

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 201)

AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Hojman

,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,

Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
R., Williamson,A., Wohlmann,P. and Wilson,R.
The Wash-Merck EST Project
Unpublished (1995)
Contact: Wilson RK

Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: ml3 -40 forward

High quality, sequence stop: 179.

FEATURES

source

1..201

/organism="Homo sapiens"

/db_xref="GDB:369856"

/db_xref="taxon:9606"

/clone="IMAGE:280210"

/clone_id="Soares,multiple_sclerosis_2NBHMSF"

/sex="male"

/tissue_type="multiple sclerosis lesions"

/dev_stage="Age 46"

/lab_host="DH10B (ampicillin resistant)"

/note="Vector: pT73D (Pharmacia) with a modified

polylinker V-TYPE: phagemid; Site 1: Not I; Site 2: Eco RI

; 1st strand CDNA was primed with a Not I - oligo(dT)

primer [5']

TGTTACCAATCTGAGTGGAGCGCCGATTTTCTTTTCTTTT 3']

double-stranded CDNA was size selected, ligated to Eco RI

adapters (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of a modified pT73 vector

(Pharmacia). Library went through one round of

normalization to a Cot = 5. Library constructed by Bento

Soares and M. Fatima Bonaldo. RNA from 4 multiple sclerosis

lesions from one patient was kindly provided by Dr. Kevin

G. Becker (NINDS/NIH).

BASE COUNT 80 a 29 c 24 g 68 t

ORIGIN

Alignment Scores:

Pred. No.: 259 Length: 201

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 12.07% Indels: 0

DB: 10 Gaps: 0

US-09-727-892a-99 (1-58) x N49179 (1-201)

Oy 46 LysLysTYRAlATYrIleIle 52

Db 112 AAAAAGTATGCTTATTATATA 132

RESULT 31

AA613897

LOCUS 208 bp mRNA linear EST 16-OCT-1997

DEFINITION no93601.s1 NCL_CGAP_P2 Homo sapiens CDNA clone IMAGE:1114392, mRNA

sequence.

ACCESSION AA613897

VERSION AA613897.1 GI:2466031

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 208)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

Tumor Gene Index

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Unpublished (1997)

JOURNAL Contact: Robert Strausberg, Ph.D.

Email: c9apbs-remail.nih.gov
 Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuagui, M.D.,
 Michael Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: David B. Krizman, Ph.D.
 CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/bbrp/image/image.html
 Insert length: 581 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 81.

FEATURES

source

Location/Qualifiers

1. .208

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="1114392"

/clone_id="NCI_CGAP_Pr2"

/sex="Male"

/dev_stage="45 years old"

/lab_host="DH10B"

/note="Vector: PAMPI0; Site_1: NotI; Site_2: EcoRI; 1st
 strand cDNA was primed with oligo(dT)17 on 50 ng of
 DNase-treated, total cellular RNA obtained from 5,000-10
 ,000 microdissected preneoplastic cells
 histologically-determined to be prostatic intraepithelial
 neoplasia 2 (PIN2) cells. Double-stranded cDNA was
 ligated to EcoRI adaptors; 5 cycles of PCR applied to the
 cDNA with an adaptor-specific primer, and the resulting
 PCR product subcloned into PAMPI0 by the DPG-cloning
 method (Life Technologies). Average insert size is 600
 bp. NOTE: Not directionally cloned. This library was
 constructed by David Krizman."

BASE COUNT

68 a 25 c 26 g 89 t

ORIGIN

Alignment Scores:

Pred. No.: 268 Length: 208
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 DB: Gaps: 0

US-09-727-892a-99 (1-58) x AA613897 (1-208)

Qy 41 Tyrasleupherlyls 47

Db 59 TATACTATTACACAAAAA 39

RESULT 32

CNS00WML

LOCUS 209 bp DNA linear GSS 28-JUN-1999

DEFINITION Arabidopsis thaliana genome survey sequence T7 end of BAC T119 of

ACCESSION TMMU library from strain Columbia of Arabidopsis thaliana, genomic

VERSION AT093771

KEYWORDS GI:5294925

SOURCE GSS.

ORGANISM Arabidopsis thaliana

REFERENCE Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:

AUTHORS Rosidae: eucosids II; Brassicales: Brassicaceae: Arabidopsis.

JOURNAL Salanoubat, M., Choisme, N., Artiguenave, F., Brothier, P., Wincker, P.,

TITLE Samsou, D., Saurin, W., Weissenbach, J. and Quetier, F.

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 209)

GENOSCOPE Genoscope.

DIRECT SUBMISSION Direct Submission

SUBMITTED (25-JUN-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr

- Web : www.genoscope.cns.fr
 Location/Qualifiers
 1. .209

/organism="Arabidopsis thaliana"

/strain="Columbia"

/db_xref="taxon:3702"

/clone_id="TMMU"

/clone_image="T119"

/note="end : T7"

BASE COUNT 53 a 45 c 37 g 74 t

ORIGIN

Alignment Scores:

Pred. No.: 269 Length: 209
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 DB: Gaps: 0

US-09-727-892a-99 (1-58) x CNS00WML (1-209)

Qy 49 AlaTyrlleleGluTyrlle 55

Db 32 GCCTATATTATAGAGTATATA 52

RESULT 33

LOCUS

DEFINITION F03788 211 bp mRNA linear EST 19-FEB-1995

ACCESSION HSC2A062 normalized infant brain CDNA Homo sapiens cDNA clone

VERSION c-2aa06 3', mRNA sequence.

KEYWORDS F03788.1 GI:670402

SOURCE EST.

ORGANISM human.

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;

Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.

1 (bases 1 to 211)

Auffray, C., Benar, G., Bois, F., Bouchier, C., da Silva, C., Devignes

, M.D., Duprat, S., Houligatte, R., Juneau, M.N., Lamy, B., Lorenzo, F.,

Mitchell, H., Mariage-Samson, R., Pletu, G., Pouliot, Y.,

Sebastiani-Kabaktchis, C. and Tessier, A.

IMAGE: molecular integration of the analysis of the human genome

and its expression

C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)

JOURNAL 95277534

MEDLINE Contact: Genethon

COMMENT Genexpress-Genethon

Genethon Centre de recherche sur le Genome Humain

1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE

Tel: 33169472800

Fax: 33160778698

Email: genexpress@genethon.fr

Single read. removed at sequence 5' end

Genexpress library idt: C; Genexpress_sequence_idt: alc-2aa06

Seq primer: (-21)M13_universal.

Location/Qualifiers

1. .211

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="c-2aa06"

/clone_id="normalized infant brain CDNA"

/sex="Female"

/tissue_type="total brain"

/dev_stage="3 months old"

/note="Organ: brain; Vector: lafmid BA; Site_1: HindIII;

Site_2: NotI; sex=Female; dev stage=3 months old;

isolate=muscular atrophy patient; tissue_type=total brain

; total mRNA was oligo-(dT) primed and directionally

cloned 5' -> 3' into the HindIII -> NotI sites of the

lafmid BA vector. Clone library from B. Soares, Psychiatry

Dept. Columbia University, USA. Normalization_method:

BASE COUNT 77 a 31 c 26 g 67 t 10 others
ORIGIN

Alignment Scores:

pred. No.: 272 Length: 211
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 10 Gaps: 0

US-09-727-892a-99 (1-58) x F03788 (1-211)

QY 46 LysLysTyrAlaTyrIleIle 52
|||||

DB 106 AAAAAGTACTTATATTATA 126

RESULT 34 222 bp DNA linear GSS 31-MAR-2000
LOCUS RPCI-23-399D1.TV RPCI-23 Mus musculus genomic clone RPCI-23-399D1,
DEFINITION
DNA sequence.

ACCESSION A2080653
VERSION A2080653.1 GI:7373554
KEYWORDS GSS.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 222)
Zhang, S., Mierman, W., Feldblum, T., Malek, J., Shatsman, S., Akintet,
B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.,
and Fraser, C.M.

TITLE Mouse BAC End Sequences from Library RPCI-23
JOURNAL unpublished (1999)
COMMENT Other-GSSs: RPCI-23-399D1.TV

CONTACT: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buhalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buhalo.edu/orderingframe.htm>)
or from Resea ch Genetics (<http://info@resgen.com>). BAC end page:
http://www.tigr.org/tadb/Dac_ends/mouse/Bac_end_intro.html
Plate: 399 row: D column: 1
Seq primer: T7
Class: BAC ends.

FEATURES
Location/Qualifiers
1..222

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-399D1"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: PBAC3.6; Site:1;
EcoRI; Site:2; EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the PBAC3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 65 a 44 c 48 g 65 t
ORIGIN

Alignment Scores:

pred. No.: 286 Length: 222
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 9 Gaps: 0

US-09-727-892a-99 (1-58) x A1681809 (1-228)

QY 43 LeupherThyLysTyrAla 49
|||||

DB 163 CTTTCACAAAATAATGCC 143

RESULT 35 228 bp mRNA linear EST 16-DEC-1999
LOCUS tx49g09.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2272960 3',
DEFINITION
mRNA sequence.

ACCESSION A1681809
VERSION A1681809.1 GI:4891991
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 228)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-f@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Sequencing by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/BLINL at:
www.dlo.llnl.gov/bdrrp/image/image.html
Insert length: 1313 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 126.

FEATURES
Location/Qualifiers
1..228

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2272960"
/clone_lib="NCI_CGAP_Lu24"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker. Plasmid DNA from the normalised
library NCI_CGAP_Lu24 was prepared, and ss circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (clonids
1414920-1417991 and 1520904-1522439). Subtraction by Bento
Soares and M. Fatima Bonaldi.

BASE COUNT 59 a 43 c 41 g 85 t
ORIGIN

Alignment Scores:

pred. No.: 294 Length: 228
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.07% Indels: 0
DB: 9 Gaps: 0

US-09-727-892a-99 (1-58) x A1681809 (1-228)

Oy 4 LysTyTyLysThrValLeu 10
 |||
 Db 25 AATACAAAACGGTCTGTG 45

RESULT 36
 F09939
 LOCUS HSC38D102 normalized infant brain cDNA Homo sapiens EST 24-FEB-1995
 DEFINITION c-38d10 3', mRNA sequence.
 ACCESSION F09939.1 GI:682472
 VERSION F09939.1
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 232)
 AUTHORS Aulfray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F., Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y., Sebastiani-Kabackchis,C. and Tessier,A.
 TITLE IMAGE: molecular integration of the analysis of the human genome and its expression
 JOURNAL C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
 MEDLINE 95277534
 COMMENT Contact: Genethon
 Genethon Centre de recherche sur le Genome Humain
 1, rue de l'Internationale, Bp60 91002 EVRY Cedex, FRANCE
 Tel: 33169472800
 Fax: 33160778698
 Email: genexpress@genethon.fr
 Single read. removed at sequence 5' end
 Genexpress_library_idt: C; Genexpress_sequence_idt: a1c-38d10
 Seq primer: (-21)M13-universal.

FEATURES
 source
 1..232
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="c-38d10"
 /clone_lib="normalized infant brain cDNA"
 /sex="Female"
 /tissue_type="total brain"
 /dev_stage="3 months old"
 /note="Organ: brain; Vector: lafmid BA; Site.1: HindIII; Site.2: NotI; sex:Female; dev_stage=3 months old; isolate-muscular atrophy patient; tissue_type=total brain; total mRNA was oligo-(dt) primed and directionally cloned 5' -> 3' into the HindIII -> NotI sites of the lafmid BA vector. Clone library from B.Soures, Psychiatry Dept. Columbia University, USA. Normalization_method: Bento Soares, P.N.A.S in press"

BASE COUNT 87 a 37 c 33 g 69 t 6 others

ORIGIN

Alignment Scores:
 Pred. No.: 299 Length: 232
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 DB: 10 Gaps: 0

US-09-727-892a-99 (1-58) x F09939 (1-232)

Oy 46 LysTyTyAlaTyrIleIle 52
 |||
 Db 57 AAAAGTATGCTTATATATA 77

RESULT 37
 F04344
 LOCUS HSC20D072 normalized infant brain cDNA Homo sapiens EST 19-FEB-1995
 DEFINITION

c-2qd07 3', mRNA sequence.
 F04344
 VERSION F04344.1 GI:670998
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 234)
 AUTHORS Aulfray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F., Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y., Sebastiani-Kabackchis,C. and Tessier,A.
 TITLE IMAGE: molecular integration of the analysis of the human genome and its expression
 JOURNAL C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
 MEDLINE 95277534
 COMMENT Contact: Genethon
 Genethon Centre de recherche sur le Genome Humain
 1, rue de l'Internationale, Bp60 91002 EVRY Cedex, FRANCE
 Tel: 33169472800
 Fax: 33160778698
 Email: genexpress@genethon.fr
 Single read. removed at sequence 5' end
 Genexpress_library_idt: C; Genexpress_sequence_idt: a1c-2qd07
 Seq primer: (-21)M13-universal.

FEATURES
 source
 1..234
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="c-2qd07"
 /clone_lib="normalized infant brain cDNA"
 /sex="Female"
 /tissue_type="total brain"
 /dev_stage="3 months old"
 /note="Organ: brain; Vector: lafmid BA; Site.1: HindIII; Site.2: NotI; sex:Female; dev_stage=3 months old; isolate-muscular atrophy patient; tissue_type=total brain; total mRNA was oligo-(dt) primed and directionally cloned 5' -> 3' into the HindIII -> NotI sites of the lafmid BA vector. Clone library from B.Soures, Psychiatry Dept. Columbia University, USA. Normalization_method: Bento Soares, P.N.A.S in press"

BASE COUNT 94 a 36 c 28 g 69 t 7 others

ORIGIN

Alignment Scores:
 Pred. No.: 302 Length: 234
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 DB: 10 Gaps: 0

US-09-727-892a-99 (1-58) x F04344 (1-234)

Oy 46 LysTyTyAlaTyrIleIle 52
 |||
 Db 99 AAAAGTATGCTTATATATA 119

RESULT 38
 A1766865
 LOCUS w189g10.x1 NCI_CGAP_K1d12 Homo sapiens cDNA IMAGE:2400546 3',
 DEFINITION mRNA sequence.
 ACCESSION A1766865
 VERSION A1766865.1 GI:5233374
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 248)
 NCBI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITL Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgap@remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40up from Gibco
 High quality sequence stop: 222.
 Location/Qualifiers
 1. 248
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2400546"
 /clone_lib="NCI-CCAP_Kid12"
 /tissue_type="2 pooled tumors (clear cell type)"
 /lab_host="DH10B"
 /note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
 a modified polylinker. Site 1: Not I; Site 2: Eco RI;
 Plasmid DNA from the normalized library NCI-CCAP_Kid5 was
 prepared, and ss circles were made in vitro. Following HAP
 purification, this DNA was used as tracer in a subtractive
 hybridization reaction. The driver was PCR-amplified cDNAs
 from a pool of 5,000 clones made from the same library
 (cloneIDs 1323912-1325831, 1471368-1472903 and
 1492104-1493255). Subtraction by Bento Soares and M.
 Fatima Bonaldo."
 BASE COUNT 99 a 36 c 29 g 84 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 320 Length: 248
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 Gaps: 0
 DB: 9
 US-09-727-892a-99 (1-58) x A1766865 (1-248)
 QY 46 LysLysTYrAlaTYrIleIle 52
 ||||||||||||||||||
 Db 105 AAAAAGTATGCTTATATATATA 125
 RESULT 39
 F03130 250 bp mRNA linear EST 02-FEB-1995
 LOCUS HSC1MB112 normalized infant brain cDNA Homo sapiens CDNA clone
 DEFINITION c-1mb11 3', mRNA sequence.
 ACCESSION F03130
 VERSION F03130.1 GI:646687
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 250)
 AUTHORS Aifray,C., Behar,G., Bols,F., Bouchier,C., da Silva,C., Devignes
 ,M.D., Duprat,S., Houligatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,
 Mitchell,H., Mariage-Samson,R., Pletcu,G., Pouliot,Y.,
 Sebastiani-Kabakchis,C. and Tessier,A.
 TITLE IMAGE: molecular integration of the analysis of the human genome
 and its expression
 JOURNAL C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
 MEDLINE 95277534

COMMENT Contact: Genethon
 Genexpress-Genethon
 Genethon Centre de recherche sur le Genome Humain
 1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
 Tel: 33169472800
 Fax: 33160778698
 Email: genexpress@genethon.fr
 Single read. removed at sequence 5' end
 Genexpress library_idt: C: Genexpress_sequence_idt: alc-1mb11
 Seq primer: (-21)M13_universal.
 Location/Qualifiers
 1. 250
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="c-1mb11"
 /clone_lib="normalized infant brain cDNA"
 /sex="Female"
 /tissue_type="total brain"
 /dev_stage="3 months old"
 /note="Organ: brain; Vector: lafmid BA; Site 1: HindIII;
 Site 2: NotI; sex=Female; dev_stage=3 months old;
 isolate=muscular atrophy patient; tissue_type=total brain
 ; total mRNA was oligo-(dT) primed and directionally
 cloned 5' -> 3' into the HindIII -> NotI sites of the
 lafmid BA vector. Clone library from B. Soares, Psychiatry
 Dept. Columbia University, USA. Normalization method:
 Bento Soares, P.N.A.S in press"
 BASE COUNT 96 a 38 c 33 g 73 t 10 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 323 Length: 250
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.07% Indels: 0
 Gaps: 0
 DB: 10
 US-09-727-892a-99 (1-58) x F03130 (1-250)
 QY 46 LysLysTYrAlaTYrIleIle 52
 ||||||||||||||||||
 Db 121 AAAAAGTATGCTTATATATATA 141
 RESULT 40
 AA353041/c
 LOCUS AA353041 256 bp mRNA linear EST 21-APR-1997
 DEFINITION EST61195 Activated T-cells XX Homo sapiens CDNA 5' end similar to
 AA353041
 ACCESSION AA353041
 VERSION AA353041.1 GI:2005432
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 256)
 AUTHORS Adams,M.D., Soares,M.B., Kerlavage,A.R., Fields,C. and Venter,J.C.
 TITLE Rapid cDNA sequencing (expressed sequence tags) from a
 directionally cloned human infant brain cDNA library
 JOURNAL Nature Genet. 4, 373-380 (1993)
 MEDLINE 94004965
 COMMENT Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3016699056
 Fax: 3016699423
 Email: arkerlavet@igf.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.


```

FEATURES
source
    Location/Qualifiers
        1..256
            /organism="Homo sapiens"
            /db_xref="ATCC (Inhost):153320"
            /db_xref="taxon:9606"
            /clone_lib="Activated T-cells XX"
            /cell_type="T-lymphocyte"
            /dev_stage="adult"
            /note="Vector: pBluescript SK-, Site_1: EcoRI; Site_2:
            XhoI"
BASE COUNT      96 a      41 c      50 g      67 t      2 others
ORIGIN
Alignment Scores:
Pred. No.:      331      Length:      256
Score:          7.00      Matches:      7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     12.07%      Indels:      0
DB:              9      Gaps:      0

US-09-727-892a-99 (1-58) x AA353041 (1-256)

OY      6  LysTrValleuLeuTyrcys 12
LOCUS   253  AAGACGGCTCTGCTATATATGC 233
DEFINITION
mRNA sequence.
ACCESSION  AA732601      264 bp      mRNA      linear      EST 23-JAN-1998
LOCUS      AA732601      n285c03.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1302244 3',
VERSION     AA732601.1  GI:2753208
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 264)
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
JOURNAL    Contact: Robert Strausberg, Ph.D.
            Email: cgapsb@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/ILNLR at:
            www-bio.lnlnl.gov/bhrp/image/image.html
            Insert length: 1160      Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 187.
            Location/Qualifiers
                1..264
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone_lib="NCI_CGAP_GCB1"
                    /clone_lib="NCI_CGAP_GCB1"
                    /tissue_type="germinal center B cell"
                    /lab_host="DH10B"
                    /note="Vector: p773D-Pac (Pharmacia) with a modified
                    polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                    was prepared from human tonsillar cells enriched for
                    germinal center B cells by flow sorting (CD20+, IgD-),
                    provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                    (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
                    primed with a Not I - oligo(dT) primer

```

```

[5'-TGTACCATCTGAGTGGAGCGCCGCCCTATTTTTTTTTTTTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      96 a      39 c      31 g      98 t
ORIGIN
Alignment Scores:
Pred. No.:      341      Length:      264
Score:          7.00      Matches:      7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     12.07%      Indels:      0
DB:              9      Gaps:      0

US-09-727-892a-99 (1-58) x AA732601 (1-264)

OY      46  LysLysTYrAlaTyrllelle 52
LOCUS   118  AAAAAGTATGCTTATATATA 138
DEFINITION
BMA24468      267 bp      mRNA      linear      EST 29-JAN-2002
LOCUS      BMA24468/c      Ipsp00202 Spleen cDNA library Ictalurus punctatus cDNA 5', mRNA
sequence.
ACCESSION  BMA24468      BMA24468
VERSION     BMA24468.1  GI:18392992
KEYWORDS    EST.
SOURCE      channel catfish.
ORGANISM    Ictalurus punctatus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Siluriformes;
            Ictaluridae; Ictalurus.
            1 (bases 1 to 267)
REFERENCE   1
            Kocabas,A., Li,P., Cao,D., Ju,Z., Karst,A., Patterson,A., Dunham,R.
            and Liu,Z.
            Transcriptome of Channel Catfish: Analysis of Expressed Sequence
            Tags, Expression Profiles, and Markers from the Spleen
            Unpublished (2002)
JOURNAL    Contact: Liu ZJ
            The Fish Molecular Genetics and Biotechnology Laboratory,
            Department of Fisheries and Allied Aquacultures and Program of Cell
            and Molecular Biosciences
            Auburn University
            203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
            Tel: 334 844 4054
            Fax: 334 844 9208
            Email: zliu@acesag.auburn.edu
            Seq primer: M13 Reverse.
            Location/Qualifiers
                1..267
                    /organism="Ictalurus punctatus"
                    /db_xref="taxon:7998"
                    /clone_lib="Spleen cDNA library"
                    /note="Organ: Spleen; Vector: pSPori1; Site_1: NotI;
                    Site_2: SalI"
FEATURES
source
    Location/Qualifiers
        1..267
            /organism="Ictalurus punctatus"
            /db_xref="taxon:7998"
            /clone_lib="Spleen cDNA library"
            /note="Organ: Spleen; Vector: pSPori1; Site_1: NotI;
            Site_2: SalI"
BASE COUNT      105 a      27 c      34 g      101 t
ORIGIN
Alignment Scores:
Pred. No.:      345      Length:      267
Score:          7.00      Matches:      7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     12.07%      Indels:      0
DB:              10      Gaps:      0

US-09-727-892a-99 (1-58) x BMA24468 (1-267)

OY      4  LysTYrLysTrValleuLeu 10

```

```

|||||
Db      179  AAATCAAAACGTTTGTCTG 159

RESULT 43
BB287799/c
LOCUS   BB287799          270 bp    mRNA    linear    EST 09-JUL-2000
DEFINITION BB287799, RIKEN full-length enriched, 2 cells egg Mus musculus cDNA
            clone BB20022D01 3', mRNA sequence.
ACCESSION BB287799
VERSION   BB287799.1  GI:8988248
KEYWORDS  EST.
SOURCE    house mouse.
           Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
           1 (bases 1 to 270)
AUTHORS  Kono,H., Alzawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
           ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
           Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
           Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
           Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.,
           Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
           Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
           ,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugihara,Y.,
           Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toya
           ,T., Tsunoda,Y., Watabiki,A., Watanabe,S., Yamamura,T., Yamanaoka,I.,
           Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
           ,M., Muramatsu,M. and Hayashizaki,Y.
           RIKEN Mouse ESTs (Kono,H., et al.)
           Unpublished (2000)
TITLE     Contact: Yoshihide Hayashizaki
           Laboratory for Genome Exploration Research Group, RIKEN Genomic
           Sciences Center(GSC), Yokohama Institute
           The Institute of Physical and Chemical Research (RIKEN)
           1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
           Tel: 81-45-503-9222
           Fax: 81-45-503-9216
           Email: genome-res@gsc.riken.go.jp,
           URL: http://genome.gsc.riken.go.jp/
           Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
           ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
           Thermostabilization and thermoactivation of thermostable enzymes by
           trehalose and its application for the synthesis of full length
           cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
           Itoh,M., Kitsuai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
           Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
           ,Y. and Hayashizaki,Y.
           Automated filtration-based high-throughput plasmid preparation
           system. Genome Res. 9 (5), 463-470 (1999)
           Carninci,P. and Hayashizaki,Y.
           High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
           19-44 (1999)
           Please visit our web site (http://genome.rtc.riken.go.jp) for
           further details.
FEATURES
SOURCE    Location/Qualifiers
           1..270
           /organism="Mus musculus"
           /db_xref="taxon:10090"
           /clone="BB20022D01"
           /clone_1lb="RIKEN full-length enriched, 2 cells egg"
           /tissue_type="egg"
           /dev_stage="2 cells"
           /lab_host="DH10B"
           /note="Site 1: SalI; Site 2: BamHI; cDNA library was
           prepared and sequenced in Mouse Genome Encyclopedia
           Project of Genome Exploration Research Group in Riken
           Genomic Sciences Center and Genome Science Laboratory in
           RIKEN. Division of Experimental Animal Research in Riken
           contributed to prepare mouse tissues. 1st strand cDNA was
           primed with a primer [5'
           GAGGAGAGAGAGATCCAGAGCTTTTCTTTTCTTTTCTTNN 3'], cDNA was
           prepared by using trehalose thermo-activated reverse
           transcriptase and subsequently enriched for full-length by

```

```

cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGGAGAGAGATCCAGAGTTCGAGTTAATTAATTCACCCCGCCCGCC 3']. cDNA
was cleaved with XhoI and BamHI. Vector: a modified
pBluescript KS(+) after bulk excision from lambda FIC I."
BASE COUNT      90 a      27 c      42 g      111 t
ORIGIN
Alignment Scores:
Pred. No.:      349      Length:      270
Score:          7.00      Matches:      7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      12.07%      Indels:      0
DB:              9      Gaps:      0
US-09-727-892a-99 (1-58) x BB287799 (1-270)
QY      5  TyrLysThrValLeuLeuTyr 11
Db      133  TATAAACTGTTATATCTAT 113

RESULT 44
BB287799/c
LOCUS   BB287799          272 bp    DNA    linear    GSS 27-AUG-1998
DEFINITION BB287799, RIKEN full-length enriched, 2 cells egg Mus musculus cDNA
ACCESSION BB287799
VERSION   BB287799.1  GI:8988248
KEYWORDS  EST.
SOURCE    human.
           Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
           1 (bases 1 to 272)
AUTHORS  Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
           Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
           Hood,L.
           Sequence-tagged connectors: A sequence approach to mapping and
           scanning the human genome
           Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
           99380589
           Contact: Mahairas GG, Wallace JC, Hood L
           High Throughput Sequencing Center
           University of Washington
           401 Queen Anne Avenue North, Seattle, WA 98109, USA
           Tel: (206) 616-3618
           Fax: (206) 616-3887
           Email: jwallace@u.washington.edu
           Sequence Tagged Connector
           Plate: 3044 row: I column: 7
           Class: BAC ends
           High quality sequence stop: 272.
FEATURES
SOURCE    Location/Qualifiers
           1..272
           /organism="Homo sapiens"
           /db_xref="taxon:9606"
           /clone="plate=3044 Col=7 Row=I"
           /clone_1lb="CIT Approved Human Genomic Sperm Library D"
           /sex="male"
           /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
           E-coli DH10B"
BASE COUNT      59 a      60 c      42 g      111 t
ORIGIN
Alignment Scores:
Pred. No.:      351      Length:      272
Score:          7.00      Matches:      7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      12.07%      Indels:      0
DB:              12      Gaps:      0

```

US-09-727-892a-99 (1-58) x AA098659 (1-272)

OY 6 LysThrValLeuLeuTyrCys 12
 |||||||
 Db 59 AAACAGTCTTTTGTACTGT 79

RESULT 45
 AA975953/c

LOCUS 275 bp mRNA linear EST 22-MAY-1998
 DEFINITION 0034b03.s1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1568045 3',
 mRNA sequence.

ACCESSION AA975953
 VERSION AA975953.1 GI:3151745

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 275)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index.

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/bdrrp/image/image.html

Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 208.

FEATURES
 source

1..275
 location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1568045"

/clone_id="NCI_CGAP_Lu5"

/tissue_type="carcinoid"

/lab_host="DH10B"

/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
 modified polylinker; 1st strand cDNA was prepared from
 neuroendocrine lung carcinoid, and was then primed with a
 Not I - oligo(dT) primer. Double-stranded cDNA was ligated
 to Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT73 vector. Library is normalized. Library was
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 124 a 41 c 33 g 77 t
 ORIGIN

Alignment Scores:

Pred. No.: 355 Length: 275

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 12.07% Indels: 0

DB: 9 Gaps: 0

US-09-727-892a-99 (1-58) x AA975953 (1-275)

OY 41 TyrAsnLeuPheThrLysLys 47
 |||||||

Db 23 TATACTATTCACAAAAA 3

Search completed: November 5, 2002, 04:53:03
 Job time : 1741 secs

